

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 17:12:23 ; Search time 11580.7 Seconds
(without alignments)
18610.264 Million cell updates/sec

Title: US-10-782-244-1

Perfect score: 5662

Sequence: 1 agcgggtgtgactgaacc.....gctggtattagagagggaac 5662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5342.4	94.4	7311	3	CR749280 Homo sapi
2	2725.4	48.1	4696	3	AL834497 Homo sapi
3	2146.6	37.9	2497	3	AK036149 Mus muscu
4	2139.8	37.8	3881	3	AK082586 Mus muscu
5	1047.6	18.5	1475	9	AY417033 Homo sapi
6	1001.6	17.7	1977	3	AK075662 Mus muscu
7	880.8	15.6	1453	9	AY417035 Mus muscu
8	830.4	14.7	1211	3	AK087858 Mus muscu
9	824.8	14.6	872	5	BQ049342 AGENCOURT
10	755.2	13.3	856	4	BG232200 602648109
11	753.2	13.3	878	6	CD653846 AGENCOURT
12	728.8	12.9	831	4	BI601569 603250782
13	721.6	12.7	747	7	CR629809 DKEZp469N
14	710.2	12.5	867	4	BG414183 GK22-237
15	704	12.4	4697	3	AK029196 Mus muscu
16	702.6	12.4	893	5	BU746120 CH3#002_A
17	700	12.4	3571	3	AK019537 Mus muscu
18	693.8	12.3	1377	9	AY417034 Pan trogl
19	690.8	12.2	801	7	CF723187 UI-M-GV0-
20	690.4	12.2	1632	3	AK054119 Mus muscu
21	680	12.0	807	5	BU751233 CH3#038 F
22	646.4	11.4	846	4	BG573617 602594635
23	640.8	11.3	645	7	CV024787 2284 Full
24	638.4	11.3	822	5	BU746121 CH3#002_A

25	636.8	11.2	676	4	BI463004
c	627.8	11.1	807	5	BU751232
27	625	11.0	825	6	CD657315
c	617.2	10.9	674	5	AX667705
29	612.6	10.8	659	5	AX666671
30	608.6	10.7	748	7	CNS34331
31	599.8	10.6	791	7	CK637702
32	595.2	10.5	841	7	CNS35250
33	589.6	10.4	664	2	BB613723
34	587	10.4	835	6	CD655466
35	585.4	10.3	651	6	BY722545
36	584	10.3	700	1	AU253730
37	579	10.2	671	6	CB434014
38	569.4	10.1	582	5	BP314060
39	563	9.9	581	5	BP314047
40	563	9.9	640	6	BY740237
41	562.4	9.9	581	5	BP352173
42	562	9.9	582	5	BP352842
43	560	9.9	581	5	BP199668
44	551.6	9.7	615	2	BB666805
45	550.8	9.7	703	2	BB650185

ALIGNMENTS

RESULT 1
CR749280

LOCUS CR749280 7311 bp mRNA linear HTC 19-AUG-2004

DEFINITION Homo sapiens mRNA; cDNA DKFZp686B11164 (from clone DKFZp686B11164).

ACCESSION CR749280

VERSION CR749280.1 GI:51476289

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 7311)

AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.

CONSRM The German CDNA Consortium

TITLE Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764

JOURNAL Neuherberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686B11164) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering. http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686B11164 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..7311
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686B11164"
/tissue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiI"
/dev_stage="adult"
/note="hypothetical protein, differentially spliced"
1..7311
/gene="DKFZp686B11164"
866..5023
/gene="DKFZp686B11164"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18135.1"

/db_xref="GI:51476290"									
/translation="MGIIATFRSWAGIINLCKPGNSGTQSLIGVLCPINMEIRRGILL VLYDIFRLPLPVPVTEEFTEALLSDVGRFQDSWRLSDGFVAAEAKTILPHRARSPL MNYLALILSAPIRNGLEGLVEVITNSDDHISVRATILIGELHMANITLPHSHSH LHCLPTLMMASFOIPEKRLRASAALNCLKRHEMKRGPKPYSHLHDIIOKATA THOKRDYLRVQKDIPIILKDTBEALLINRDSQVLOKHENLGNWNLLIGTFLESEEDG LNYKDEQHLREVRLLLYFPKSKLYANLDDLFAKQLTIVGQCQFTEFLSEEDG QCYLEDVNDIOWMLNASGMPERSLQNGLLTLSQHYFLFTIGTUSCHPHGVKME KCSVFQCLNLCSLKNQDHLLKLTSSLDYSRDLGLARVILSKILTAAADACRLYATKH LVLLRANVEFNNGIILLVQLHDKNKTISSEALDILDEACEDKANLHALIOKPA LSHLGKGLLLRLSIPKGLSYLNERGVYAKOLEKWHREYKPYVDLIEOLNEAL TYRKPVDGNDVRRNQRLQRPVILPHLYHQVYVHKHTGCHLLEWONITELCRNV RTPDLDKWEIKKJASWALGNIGSSNWMLNLEBNVIFDLKLAQCEVLSIRGT CVTVLGLIAKTGGCDILKCHNDVARRSHKHLWFPVDDVEQLCNEUSIPSTLSLN SSTSRHNSSESVPSSMILEDDRFSGSSTSTFFLDINEDTEFTFYDRSGPTKOKN SPFPASSKLVNRIILNSILPNKKHRSDDPKGKLSSEKTSNRRIRTLITEPSVDF NHSDDTPTISTVQKTLQLETSFPGNKHLEDGTSPTSGIENDLIKFTKFTGNTENRNTS RELNVESSTSHMKIRSGSPNDDTTSGISMSKSSPSPRETVGDVATMDTDCCSMST VVSSTIKTSHYLTPOSNHLSKSSVLSVPPGSSHTLPRAQSLKAPSATIKSLA DGNFYSYTSRDAFGVATLRLQOORHPSLSHSEALSPAKVDLFTDTITMKANSFES RLTPSRFMKALSYASLDKEDLLSPINQNTLQSSVRSVSVSSATYGGSDDYIGLALPV DINDIIFOVKDIPVFOFTKNTIPPHDDRGARAFADHAGESIEDTGLQHTDDNCLYCVIE ILGFQSNOLSAICSHSDQDIPYSDWCQOTHNPLEVVPKFSGISGSDGVQSEGS ASSTKTELLGUKTIPDPTPMCRILLRVELRLVINLSSVSTKCHETGLLTKEKY PQTFDDICLYSEVSHLSHCTFLPCRRFIQELFQDVQFLOMHEAEAVLATPPKQPI VDTSAES"									
ORIGIN									
Query Match	94.4%	Score	5342.4;	DB	3;	Length	7311;		
Best Local Similarity	97.7%	Pred. No.	0;						
Matches	5518;	Conservative	0;	Mismatches	11;	Indels	119;	Gaps	5;
QY	16	AAACCCGTCATATGGCGCGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGA	75						
DB	1	AAACCCGTCATATGGCGCGGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGA	60						
QY	76	GGCGGGAATGACAGCGCGAGGAGAACGTCCTCGCTGGATCTGACCGAGAACCTTCTGAT	135						
DB	61	GGCGGGAATGACAGCGCGAGGAGAACGTCCTCGCTGGATCTGACCGAGAACCTTCTGAT	120						
QY	136	AACCTTAAGAGAGATCTCCAAAATGTGCGCAGATGTCAGGAGATGATCAAAATATGAGAAAG	195						
DB	121	AACCTTAAGAGAGATCTCCAAAATGTGCGCAGATGTCAGGAGATGATCAAAATATGAGAAAG	180						
QY	196	CTAGGCCATCTGAATAACTTTTACTAAGCTTCTTGGATATGTCGACCGAGAACCTTCTGAT	255						
DB	181	CTAGGCCATCTGAATAACTTTTACTAAGCTTCTTGGATATGTCGACCGAGAACCTTCTGAT	240						
QY	256	CTGGGCTTCACTATGAGGATATCATATTTGTTGGGTTAGCTTTTATTAATGAGCA	315						
DB	241	CTGGGCTTCACTATGAGGATATCATATTTGTTGGGTTAGCTTTTATTAATGAGCA	300						
QY	316	AAAGAAAGTGCAGCAGCAGCGCTACGAGCGCTTCGATATCTCATCCAAAGATCCAGTATT	375						
DB	301	AAAGAAAGTGCAGCAGCAGCGCTACGAGCGCTTCGATATCTCATCCAAAGATCCAGTATT	360						
QY	376	CTCCAGAAAGTGCTAAATTTGAAGTGAATTTAATAGTAGTGTGCTTGAATGACATACAA	435						
DB	361	CTCCAGAAAGTGCTAAATTTGAAGTGAATTTAATAGTAGTGTGCTTGAATGACATACAA	420						
QY	436	CAGACCAACGAGTAGAGAGACACAGCACTTCGATTAGTCAGAAAGATGATTACTGTG	495						
DB	421	CAGACCAACGAGTAGAGAGACACAGCACTTCGATTAGTCAGAAAGATGATTACTGTG	480						
QY	496	AATGCTTCTCTGTTTCTTAGTTCGTGACCACTCATTAATTTGAGTTGGAATGATGGA	555						
DB	481	AATGCTTCTCTGTTTCTTAGTTCGTGACCACTCATTAATTTGAGTTGGAATGATGGA	540						
QY	556	CTTCAAGAAAGACAGAGATGTCGAGCATGCTTGCATTTATCTGTGAACCTAGCACTT	615						
DB	541	CTTCAAGAAAGACAGAGATGTCGAGCATGCTTGCATTTATCTGTGAACCTAGCACTT	600						
QY	616	CAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAACACCATATGGAATGTGATT	675						

DB	601	CAGAAATCCAGAGGTGGTGGCCCTTCGAGTGGACTAAACACCACTCTTGAATAATGTGATC	660
QY	676	GATTGCGAATAAGTGCAGATAAATAGAGGCCCTTAATTACTACAAATTTTGACCTCTTAAT	735
DB	661	GATTGCGAATAAGTGCAGATAAATAGAGGCCCTTAATTACTACAAATTTTGACCTCTTAAT	720
QY	736	CATCCAAAGACTCGCGAGTATGTCGAGCTGATGTAGAAATTAGAGAAATTTTAGCACCC	795
DB	721	CATCCAAAGACTCGCGAGTATGTCGAGCTGATGTAGAAATTAGAGAAATTTTAGCACCC	778
QY	796	TATATCTGATTTTCACTACAGACATAGTCCAGATACAGCTGAAGACAGCTCAAGAGAC	855
DB	779	TATATCTGATTTTCACTACAGACATAGTCCAGATACAGCTGAAGACAGCTCAAGAGAC	838
QY	856	AGAGAGACAGATTTTAGCCAGTAAATGGAATCATAGCAATTTCCCATCATGGGCA	915
DB	839	AGAGAGACAGATTTTAGCCAGTAAATGGAATCATAGCAATTTCCCATCATGGGCA	898
QY	916	GGTATTATTAATTTATGTAACCTGGAAATTTCTGGGATCCAGTCTCTAATAGGATGATCTT	975
DB	899	GGTATTATTAATTTATGTAACCTGGAAATTTCTGGGATCCAGTCTCTAATAGGATGATCTT	958
QY	976	TGCATACCAATATGGAATAAGCGGAGGCTTACTTGAAGTCTTTATGATATATTTTCGT	1035
DB	959	TGCATACCAATATGGAATAAGCGGAGGCTTACTTGAAGTCTTTATGATATATTTTCGT	1018
QY	1036	CTTCTCTTACCTGTTGTGACTGAGGAGTTCATAGAAGCACTACTCAGTGTAGATCCAGGG	1095
DB	1019	CTTCTCTTACCTGTTGTGACTGAGGAGTTCATAGAAGCACTACTCAGTGTAGATCCAGGG	1078
QY	1096	AGTTTCCAAAGACAGTTGGAGGCTTTTCAGATGGCTTTTGTGAGCTGAGGCAAACTATT	1155
DB	1079	AGTTTCCAAAGACAGTTGGAGGCTTTTCAGATGGCTTTTGTGAGCTGAGGCAAACTATT	1138
QY	1156	CTTCTCTATGTCGAGATCCAGGCGACCTCATGATGATTAATTTTGGGCACTGATCTC	1215
DB	1139	CTTCTCTATGTCGAGATCCAGGCGACCTCATGATGATTAATTTTGGGCACTGATCTC	1198
QY	1216	TCGTGATTTATTCGTAATGGAATTTTGTAGAGGCTTCTAGTTGAAGTGAATAAACAAGTAT	1275
DB	1199	TCGTGATTTATTCGTAATGGAATTTTGTAGAGGCTTCTAGTTGAAGTGAATAAACAAGTAT	1258
QY	1276	GATCATATCTCAGTTAGAGCTACATCTCTTTTAGGAGAGCTTTTACATATGCGAAACACA	1335
DB	1259	GATCATATCTCAGTTAGAGCTACATCTCTTTTAGGAGAGCTTTTACATATGCGAAACACA	1318
QY	1336	ATTCTTCTCATTTACATAGCCATCATTTTACACTGCTTGGCCAACTTAATGAATGCT	1395
DB	1319	ATTCTTCTCATTTACATAGCCATCATTTTACACTGCTTGGCCAACTTAATGAATGCT	1378
QY	1396	GCATCTTTTGATATCCCAAGGAAAGAGACTGCGAGCGAGCTGCGAGCTTGAACCTGTTTA	1455
DB	1379	GCATCTTTTGATATCCCAAGGAAAGAGACTGCGAGCGAGCTTGAACCTGTTTA	1438
QY	1456	AAAAGCTTCCATGAATAAGAGAAAACGAGGACCTTAGCTTATAGTCTTTCATTTAGCCAC	1515
DB	1439	AAAAGCTTCCATGAATAAGAGAAAACGAGGACCTTAGCTTATAGTCTTTCATTTAGCCAC	1498
QY	1516	ATTATTTCAGAAAGCAATTCGCAACACACAGGAAACCGGATCAGTATCTCCGAGTTTCAAAA	1575
DB	1499	ATTATTTCAGAAAGCAATTCGCAACACACAGGAAACCGGATCAGTATCTCCGAGTTTCAAAA	1558
QY	1576	GATATATTTATTCCTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAA	1635
DB	1559	GATATATTTATTCCTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAA	1618
QY	1636	GTCTTCAACATAAAGAGATCTTGAATGGAATTTGAATCTTATAGGACCACTTCTTAAG	1695
DB	1619	GTCTTCAACATAAAGAGATCTTGAATGGAATTTGAATCTTATAGGACCACTTCTTAAG	1678
QY	1696	TGCGCAAAATGTAATCTAAGAAACTAATAAGATGAACAGTTACACAGGTTTGTACAGA	1755

Db 1679 TGGCCAAATGTAATCTAAGAAAATAAAGATGAAACAGTTTACACAGTTTGTACGAAGA 1738
Qy 1756 CTACTTTATTTTACAAGCCAGCAGTAAATATATATGCAACCTGGATCTGGATTTTGCC 1815
Db 1739 CTACTTTATTTTACAAGCCAGCAGTAAATATATATGCAACCTGGATCTGGATTTTGCC 1798
Qy 1816 AAGGCCAAACAGCTCACGGTTGTAGTTGCGAGTTTACAGAAATTTCTTGTGAATCTGAA 1875
Db 1799 AAGGCCAAACAGCTCACGGTTGTAGTTGCGAGTTTACAGAAATTTCTTGTGAATCTGAA 1858
Qy 1876 GAGGATGGGCAAGCTACTTAGAGATCTAGTAAGGATATTTGATGGCTCAATGCT 1935
Db 1859 GAGGATGGGCAAGCTACTTAGAGATCTAGTAAGGATATTTGATGGCTCAATGCT 1918
Qy 1936 TCATCTGGAATGAACCCGAAAGAGTCTTCAAAATAATGGTTTATTTGACCAACCTTAGT 1995
Db 1919 TCATCTGGAATGAACCCGAAAGAGTCTTCAAAATAATGGTTTATTTGACCAACCTTAGT 1978
Qy 1996 CAACACTACTTTTATTTATTTGGAACACTTTCTTGGCCACCTCATGTGAGTTAAATGCTG 2055
Db 1979 CAACACTACTTTTATTTATTTGGAACACTTTCTTGGCCACCTCATGTGAGTTAAATGCTG 2038
Qy 2056 GAAAAATGCAAGTGTATTTGAGTGTCTCTTAATCTTTGCTCTTGAATAACCAAGATCAC 2115
Db 2039 GAAAAATGCAAGTGTATTTGAGTGTCTCTTAATCTTTGCTCTTGAATAACCAAGATCAC 2098
Qy 2116 TTGCTAAAACTTACTGTTTCTAGCTTGTGACTATATAGCAGAGATGATGGCTAGAGTCAATC 2175
Db 2099 TTGCTAAAACTTACTGTTTCTAGCTTGTGACTATATAGCAGAGATGATGGCTAGAGTCAATC 2158
Qy 2176 CTTTCCAAAAATTTTAACTGCAGCTACTGATGCCCTGCAGACTCTATGCAACAAAAACATTTA 2235
Db 2159 CTTTCCAAAAATTTTAACTGCAGCTACTGATGCCCTGCAGACTCTATGCAACAAAAACATTTA 2218
Qy 2236 AGGCTATTTATGAGAGCTAATGTTGAATTTCTTTAATTTGGGGAATTTGATGTTAGTG 2295
Db 2219 AGGCTATTTATGAGAGCTAATGTTGAATTTCTTTAATTTGGGGAATTTGATGTTAGTG 2278
Qy 2296 ACCCAGCTACATGATAAAAAAACAACGATTTCTCTGAAGCTCTTGATATCTCTCGATGAA 2355
Db 2279 ACCCAGCTACATGATAAAAAAACAACGATTTCTCTGAAGCTCTTGATATCTCTCGATGAA 2338
Qy 2356 GCATGTGAAGACAAGGCCAATCTTCATGCTCTCATTCAGATGAAACAGCGTTATCCAC 2415
Db 2339 GCATGTGAAGACAAGGCCAATCTTCATGCTCTCATTCAGATGAAACAGCGTTATCCAC 2398
Qy 2416 CTTGGAGACAAGGGTTTGTCTTCTCTGCTGAGATTTCTCTCAATTCAAAAGGATTTTCC 2475
Db 2399 CTTGGAGACAAGGGTTTGTCTTCTCTGCTGAGATTTCTCTCAATTCAAAAGGACTTTCC 2458
Qy 2476 TATCTGAATGAAGAGCTTATGTAACAAACAAATTTGGAAGTGGCACAGGGAATACAAC 2535
Db 2459 TATCTGAATGAAGAGCTTATGTAACAAACAAATTTGGAAGTGGCACAGGGAATACAAC 2518
Qy 2536 TCCAAATATGTGACTTGTGATTTGAGGAACAACCTCAATGAAGCACTTACTTACCCGGAAG 2595
Db 2519 TCCAAATATGTGACTTGTGATTTGAGGAACAACCTCAATGAAGCACTTACTTACCCGGAAG 2578
Qy 2596 CTTGTTGATGGTGATATCTATGTTGCTGGAGTAACCAAGATTAAGAGTCTCTCAGTC 2655
Db 2579 CTTGTTGATGGTGATATCTATGTTGCTGGAGTAACCAAGATTAAGAGTCTCTCAGTC 2638
Qy 2656 TACTGCTTATACACCTTTATGGCAACTAGTACACCATTAACAGGCTGCCATTTGTTG 2715
Db 2639 TACTGCTTATACACCTTTATGGCAACTAGTACACCATTAACAGGCTGCCATTTGTTG 2698
Qy 2716 GAAGTACAGATATTTATTTACAGAACTCTGTCGTAATGTTGCTACACAGATTTGGATAAG 2775
Db 2699 GAAGTACAGATATTTATTTACAGAACTCTGTCGTAATGTTGCTACACAGATTTGGATAAG 2758
Qy 2776 TGGGAAGAAATTTAAAAAATGAAAGCATCTCTTTGGGCCCTTGGGAATATCGGCTCATCA 2835
Db 2759 TGGGAAGAAATTTAAAAAATGAAAGCATCTCTTTGGGCCCTTGGGAATATCGGCTCATCA 2818

Qy 2836 AATTGGGGTCTCAATTTGCTACAGGAAGAAAAGTGTATTCAGATATATAAAAACTTGCA 2895
Db 2819 AATTGGGGTCTCAATTTGCTACAGGAAGAAAAGTGTATTCAGATATATAAAAACTTGCA 2878
Qy 2896 AATCAGTGTGAAGTCTTTTCCATCAGAGGAACTGTGTATATATGTACTTGGGCTCATAGCT 2955
Db 2879 AATCAGTGTGAAGTCTTTTCCATCAGAGGAACTGTGTATATGTACTTGGGCTCATAGCT 2938
Qy 2956 AATCAGTGTGAAGTCTTTTCCATCAGAGGAACTGTGTATATATGTACTTGGGCTCATAGCT 3015
Db 2939 AATCAGTGTGAAGTCTTTTCCATCAGAGGAACTGTGTATATATGTACTTGGGCTCATAGCT 2998
Qy 3016 CGCAAAACATCTGTGGCCAGTGGTTCCAGATGATGTGAAACAACTCTGTATATGAACCTTTCA 3075
Db 2999 CGCAAAACATCTGTGGCCAGTGGTTCCAGATGATGTGAAACAACTCTGTATATGAACCTTTCA 3058
Qy 3076 TCTATCCCAAGCACTCTTAAGTTTGAACCTCGAGTCAACAGCTCTAGACATAATAGTGAA 3135
Db 3059 TCTATCCCAAGCACTCTTAAGTTTGAACCTCGAGTCAACAGCTCTAGACATAATAGTGAA 3118
Qy 3136 AGTGAATCTGTGCGCATCGAGTATGTTCAATATGGAGGATGACCGGTTTGGCAGCAGCTCT 3195
Db 3119 AGTGAATCTGTGCGCATCGAGTATGTTCAATATGGAGGATGACCGGTTTGGCAGCAGCTCT 3178
Qy 3196 ACTAGTACATTTTCTTGTATATCAATGAAGATACAGAGCAACATTTTATGACCGATCT 3255
Db 3179 ACTAGTACATTTTCTTGTATATCAATGAAGATACAGAGCAACATTTTATGACCGATCT 3238
Qy 3256 GGAACCCATAAAGGATAAAAATCAATCCCTTTCTTGTCTTCTAGTAACTTTGTGAAGAAT 3315
Db 3239 GGAACCCATAAAGGATAAAAATCAATCCCTTTCTTGTCTTCTAGTAACTTTGTGAAGAAT 3298
Qy 3316 CGTATCTTAAATTTGCTTACTTTTGCCTTAACAAAAACATCGTAGTAGCAGTGTCCAAA 3375
Db 3299 CGTATCTTAAATTTGCTTACTTTTGCCTTAACAAAAACATCGTAGTAGCAGTGTCCAAA 3358
Qy 3376 GGAAGGAAATTTATCATCTGAAAGTAAGACAAGCAACAGGCGAATCAGAACACTTTACCGAG 3435
Db 3359 GGAAGGAAATTTATCATCTGAAAGTAAGACAAGCAACAGGCGAATCAGAACACTTTACCGAG 3418
Qy 3436 CCAGTGTGATTTTAAATCATAGTATGATTTTACACCCATATCCACTGTACAGAAAACA 3495
Db 3419 CCAGTGTGATTTTAAATCATAGTATGATTTTACACCCATATCCACTGTACAGAAAACA 3478
Qy 3496 TTACAATTTAGAGACTTTCAATTTATGGGGAATAAGCAATTTGAAGACACTGTGTACACCA 3555
Db 3479 TTACAATTTAGAGACTTTCAATTTATGGGGAATAAGCAATTTGAAGACACTGTGTACACCA 3538
Qy 3556 AGCATTTGGAGAAATGCACTTTAAATTTCAACCAAGAAATTTTGGTACAGAGAATCAGAGAA 3615
Db 3539 AGCATTTGGAGAAATGCACTTTAAATTTCAACCAAGAAATTTTGGTACAGAGAATCAGAGAA 3598
Qy 3616 AATPACAGCCGAGAGAGGTTAGTAGTAAAGTTTCAACAGCTCACATATGAAGATACGT 3675
Db 3599 AATPACAGCCGAGAGAGGTTAGTAGTAAAGTTTCAACAGCTCACATATGAAGATACGT 3658
Qy 3676 AGCCAAAGTTTCAATACAGACACTACAAAGTGGCAATAGTTTCAATGAGCTCAAGTCTCT 3735
Db 3659 AGCCAAAGTTTCAATACAGACACTACAAAGTGGCAATAGTTTCAATGAGCTCAAGTCTCT 3718
Qy 3736 TCACGAGAGACACTAGTGTAGATGCTTACAACTATGACACAGACTGTGGAAGCATGAGT 3795
Db 3719 TCACGAGAGACACTAGTGTAGATGCTTACAACTATGACACAGACTGTGGAAGCATGAGT 3778
Qy 3796 ACTGTGTAAGTACTAAAACTATTAAAGCAAGCCACTATTTTGACGCCACAGTCTTAACCAT 3855
Db 3779 ACTGTGTAAGTACTAAAACTATTAAAGCAAGCCACTATTTTGACGCCACAGTCTTAACCAT 3838
Qy 3856 CTGTCTCTCTCAAATCAAATTTGGTGTCTCGGTGCTCCAGGTTCTTCTCATACGCTT 3915
Db 3839 CTGTCTCTCTCAAATCAAATTTGGTGTCTCGGTGCTCCAGGTTCTTCTCATACGCTT 3898

Qy	3916	CCTAGAGAGCAGAGTCCCTTAAAGACACCTCTATTGCTGTAACAATTAAGAGTCTAGCAGAT	3975
Db	3899	CCTAGAGAGCAGAGTCCCTTAAAGACACCTCTATTGCTGTAACAATTAAGAGTCTAGCAGAT	3958
Qy	3976	TGTAACCTTTAGTTACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAGAAGCTA	4035
Db	3959	TGTAACCTTTAGTTACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAGAAGCTA	4018
Qy	4036	CAGCAACAAAAGATTCGATCCCTATCTACCTCTCAAGCTTTGGCATCTCCAGACAAA	4095
Db	4019	CAGCAACAAAAGATTCGATCCCTATCTCACTCTGAAGCTTTGGCATCTCCAGACAAA	4078
Qy	4096	GATGTGCTATTACTGATACCATCAACATGAAGCCAAAGTTTGTAGTCCAGATTAAACA	4155
Db	4079	GATGTGCTATTACTGATACCATCAACATGAAGCCAAAGTTTGTAGTCCAGATTAAACA	4138
Qy	4156	CCAGCAGGTTCAAGAAAGCCTTAAGTTATGTCATCATTAAGTAAGAAGATTTATTGAGT	4215
Db	4139	CCAGCAGGTTCAAGAAAGCCTTAAGTTATGTCATCATTAAGTAAGAAGATTTATTGAGT	4198
Qy	4216	CCATTAAATCAAAATACCCCTGCAACGATCTTCCTCAGTGGCTCCATGCTGCAGTGCC	4275
Db	4199	CCATTAAATCAAAATACCCCTGCAACGATCTTCCTCAGTGGCTCCATGCTGCAGTGCC	4258
Qy	4276	ACATATGGGGTTCCAGATGATTAACATTTGCTCTCTCCCGGTGGATATAAATGATATA	4335
Db	4259	ACATATGGGGTTCCAGATGATTAACATTTGCTCTCTCCCGGTGGATATAAATGATATA	4318
Qy	4336	TTCCAGGTTAAGGATATTCCTATTTTCAGACAAAAACATACCAACACATGATGTCGA	4395
Db	4319	TTCCAGGTTAAGGATATTCCTATTTTCAGACAAAAACATACCAACACATGATGTCGA	4378
Qy	4396	GGTGCAAGAGCATTTGCCCATGATGACAGGAGTCTTCCATCTCGAACTGGAGGCTTGTA	4455
Db	4379	GGTGCAAGAGCATTTGCCCATGATGACAGGAGTCTTCCATCTCGAACTGGAGGCTTGTA	4407
Qy	4456	AAAAATTTCTTTTCACTTGCTACGACGACAGATGAGTCTTAGCGAAATAATGAATTCAAATC	4515
Db	4408	-----	4407
Qy	4516	CATTCAGATGCCTCTCTGTTTTTTAGAAAGTACAGAGACACTGGACTACAGGAACATACA	4575
Db	4408	-----AGAAAGTACAGAGACACTGGACTACAGGAACATACA	4444
Qy	4576	GATGATAACTGCCCTTTATGTGCTGATTTGAATCTCGGTTTCCAGCCCGACCAACAA	4635
Db	4445	GATGATAACTGCCCTTTATGTGCTGATTTGAATCTCGGTTTCCAGCCCGACCAACAA	4504
Qy	4636	CTGAGTGCAATATGTAGTCAATTCAGACATTTCCATATCTGATTTGGTGTGAG	4695
Db	4505	CTGAGTGCAATATGTAGTCAATTCAGACATTTCCATATCTGATTTGGTGTGAG	4564
Qy	4696	CAGACTATCCATAATCCTTTAGAGTGGTTCCTCTAAGTTTTCGGGGAATTTCTGGAATGC	4755
Db	4565	CAGACTATCCATAATCCTTTAGAGTGGTTCCTCTAAGTTTTCGGGGAATTTCTGGAATGC	4624
Qy	4756	AGTGATGGGTGTCTCAAGAGGCTCAGTAGCAGACCAAAAGCACAGAAATTTGTACTA	4815
Db	4625	AGTGATGGGTGTCTCAAGAGGCTCAGTAGCAGACCAAAAGCACAGAAATTTGTACTA	4684
Qy	4816	GGTGTTAAAAACAATTCAGATGATACACCAATGTGCCGTATATCTCTTCGCAAGAAGTT	4875
Db	4685	GGTGTTAAAAACAATTCAGATGATACACCAATGTGCCGTATATCTCTTCGCAAGAAGTT	4744
Qy	4876	CTAAGATTAGTCATTAATTTGAGTGTCAAGTTTCAACTAAATGTCAATGAGACTGGGCTT	4935
Db	4745	CTAAGATTAGTCATTAATTTGAGTGTCAAGTTTCAACTAAATGTCAATGAGACTGGGCTT	4804
Qy	4936	TTAACCAATTAAGGAGAGTATCCTCAACATTTGATGACATATGCCCTTTACTCTGAGGTT	4995
Db	4805	TTAACCAATTAAGGAGAGTATCCTCAACATTTGATGACATATGCCCTTTACTCTGAGGTT	4864
Qy	4996	TCCCAATTTGCTGTCCACTGCAATTCAGACTTTCCGTGTGGAGGTTTCAACAGAATTA	5055

[illegible]

Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761C062>
 Further information about the clone and the sequencing project is
 available at <http://mips.gsf.de/projects/cdna/>.

```

FEATURES
    source            1..4696
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="RZPD:DKFZp761C062"
                    /db_xref="taxon:9606"
                    /clone="DKFZp761C062"
                    /tissue_type="amygdala"
                    /clone_lib="761 (synonym: hamy2). Vector pSport1; host
                    DH10B; sites NotI + SalI"
                    /dev_stage="adult"
                    /note="hypothetical protein, N-terminus truncated,
                    differentially spliced"
    gene             1..4696
                    /gene="DKFZp761C062"
    CDS               <3..2408
                    /gene="DKFZp761C062"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="CAD39155.1"
                    /db_xref="GI:21740291"
                    /db_xref="UniProt/TREMBL:Q8N3A0"
                    /translation="NIGSNWGLNLQEEVIVPDLILKAKQCEVLSIRGTCVVVLGLI
                    AKTKQGGDILKCHNWDVRSRHLWPVDDVEOLCNELSPISLINSSESSRH
                    NSSEVPSSPFLIEDDFGSSSTPFELDINEDTEFPYDRSGPIKDKNSPPFPASS
                    KLVNRILNSLTLPNKHRSRSPDKGKLSSEKTSNRRILRTFESVDFNDDPTP
                    ISTVQTLQLETSPFMGNKHIEDTGSPSIGENDLKFNFNGFNHRENTSRELRVSES
                    STSHMKIRSQSNTDTTSGISMSSPSRETGVGDATTMDTDCGSMSTVSVTKTIK
                    TSHYLTTPQSNHLSLSKNSVSLVPPGSSHTLPRAOSLKAPSIATIKSLADCNFSVTS
                    SRDAFGVATLKRQOORMPSLSHSBAASPADKVLFTDTITMKANSFESRLTPSRID
                    FKXKHVGISLRPTLTNLFPMKALSVASLDKEDLLSPINQNTLQRSSVSRSVSS
                    ATYGGSDYIGLALPVDINDI FQVKDIPYQTKNIPPHDRGARAFHADAGGLPSGTG
                    GLVNSFHLLRQMSLTIENNS IHSDSLFELEDTEGLQEHDDNCLYCVCEIILGF
                    QPNQSLGACSHSDFOPIYSDWCEQTTHNLPVPSKFSGISGSDGVSQEGSASST
                    KSTELLIGVKTIPDDTPMCRILLRKEVLRVLNLSVSVTKHETGLLTIKKYPOTF
                    DDICLYSEVSHLLSHCTFRLPCRFRFQELFQDVFQFLQMBEAEAVLATPBPQPIVDTs
                    AES"
    ORIGIN
    Query Match      48.1%; Score 2725.4; DB 3; Length 4696;
    Best Local Similarity 97.4%; Pred. No. 0;
    Matches 2841; Conservative 0; Mismatches 1; Indels 75; Gaps 4;

    QY 2819 GAAATATCGGCTCATCAAAATGGGGTCTCAATTTGCTACAGGAGAAACGTGATCCAG 2878
    DB 1 GAAATATCGGCTCATCAAAATGGGGTCTCAATTTGCTACAGGAGAAACGTGATCCAG 60

    QY 2879 ATATACATAAACTTGCACAAACAGTGTGAAGTCTTCTCCATCAGAGGACCTGTGTATATG 2938
    DB 61 ATATACATAAACTTGCACAAACAGTGTGAAGTCTTCTCCATCAGAGGACCTGTGTATATG 120

    QY 2939 TACTTGGGCTCATAGCTATAAACCAAAACAAGCGTGTGATATTTCTAAATGTGCACAACTGGG 2998
    DB 121 TACTTGGGCTCATAGCTATAAACCAAAACAAGCGTGTGATATTTCTAAATGTGCACAACTGGG 180

    QY 2999 ATGCTGTAGGCGATAGTCGCAAAACATCTGTGGCAGTGGTTCCAGATGATGTGGAACAAC 3058
    DB 181 ATGCTGTAGGCGATAGTCGCAAAACATCTGTGGCAGTGGTTCCAGATGATGTGGAACAAC 240

    QY 3059 TCTGTAATGAACCTTTCATCTATCCACAGCACTCTTAAGTTTGAATCGGAGTCAACAGCT 3118
    DB 241 TCTGTAATGAACCTTTCATCTATCCACAGCACTCTTAAGTTTGAATCGGAGTCAACAGCT 300

    QY 3119 CTAGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTCGAGGATGACC 3178
    DB 301 CTAGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTCGAGGATGACC 360

    QY 3179 GTTTGGCAGCAGCTCTACCTAGTACATTTTCTTGTATCAATCAATGAAGATACAGAGCCAA 3238
    DB
  
```

```

    DB 361 GGTGTGGCAGCAGCTCTACTAGTAGTACATTTTTCTTGATATCAATGAAGATACAGAGCCAA 420
    QY 3239 CATTTATAGCCAGTCTGGACCCATAAAGGATAAAAATTCATTCCTCTTCTTCTTCTCTTA 3298
    DB 421 CATTTATAGCCAGTCTGGACCCATAAAGGATAAAAATTCATTCCTCTTCTTCTTCTTCTTA 480
    QY 3299 GTAAACTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCTTAAACAAAAACATCGTA 3358
    DB 481 GTAAACTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCTTAAACAAAAACATCGTA 540
    QY 3359 GTAGCAGTGTCCAAAGAGGAGGAAATATCATCTGAAGTAAAGCAAGCAAGCAAGGCGAA 3418
    DB 541 GTAGCAGTGTCCAAAGAGGAGGAAATATCATCTGAAGTAAAGCAAGCAAGCAAGGCGAA 600
    QY 3419 TCAGAACACATCTAGCGAGCCAGTGTGATTTAAATTCGCTTACTTTGCTTAAACAAAAACATCGTA 3478
    DB 601 TCAGAACACATCTAGCGAGCCAGTGTGATTTAAATTCGCTTACTTTGCTTAAACAAAAACAT 660
    QY 3479 CCACCTGTACAGAAACATTAACAATTAAGAGACTTCAATTTATGGGGAATTAAGCACAATTGAAG 3538
    DB 661 CCACCTGTACAGAAACATTAACAATTAAGAGACTTCAATTTATGGGGAATTAAGCACAATTGAAG 720
    QY 3539 ACACCTGTAGTACACCAAGCAATTTGGAGAAAATGACTTTAAATTCACCAAGAAATTTGGTA 3598
    DB 721 ACACCTGTAGTACACCAAGCAATTTGGAGAAAATGACTTTAAATTCACCAAGAAATTTGGTA 780
    QY 3599 CAGAGAAATCACAGAGAAATACAAAGCCGAGAGAGGTTAGTAGTAAGTTCACACGAGCT 3658
    DB 781 CAGAGAAATCACAGAGAAATACAAAGCCGAGAGAGGTTAGTAGTAAGTTCACACGAGCT 840
    QY 3659 CACATATGAAGATACGTAGCCAAAGTTTCAATTAAGAGACTTCAATTAAGAGGAGTTCACAGT 3718
    DB 841 CACATATGAAGATACGTAGCCAAAGTTTCAATTAAGAGACTTCAATTAAGAGGAGTTCACAGT 900
    QY 3719 CAATGAGCTCAAGTCTCTTACAGAGAGAGAGTGTAGTAGTCAACTATGACACAG 3778
    DB 901 CAATGAGCTCAAGTCTCTTACAGAGAGAGAGTGTAGTAGTCAACTATGACACAG 960
    QY 3779 ACTGTGGAAGCATGAGTACTGTGTAAGTAACTAACTATTAAGCAAGCAAGTTCATTTGA 3838
    DB 961 ACTGTGGAAGCATGAGTACTGTGTAAGTAACTAACTATTAAGCAAGCAAGTTCATTTGA 1020
    QY 3839 CGCCACAGTCTAAACATCTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTCTCCAG 3898
    DB 1021 CGCCACAGTCTAAACATCTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTCTCCAG 1080
    QY 3899 GTTCTTCTCATAGCTTCTTACAGAGAGAGTCCCTTAAAGCAAGTCTTATTTGCTACAA 3958
    DB 1081 GTTCTTCTCATAGCTTCTTACAGAGAGAGTCCCTTAAAGCAAGTCTTATTTGCTACAA 1140
    QY 3959 TTAAGAGTCTAGCAGATTTGTAATTTAGTAACTTACAGAGTCTAGAGATGCTTTGGCTATG 4018
    DB 1141 TTAAGAGTCTAGCAGATTTGTAATTTAGTAACTTACAGAGTCTAGAGATGCTTTGGCTATG 1200
    QY 4019 CTACACTGAAAGACTACAGCAACAAAGATTCATCTTATCTTCTCTCTCTCTCTCTCTCTCT 4078
    DB 1201 CTACACTGAAAGACTACAGCAACAAAGATTCATCTTATCTTCTCTCTCTCTCTCTCTCTCT 1260
    QY 4079 TGGCATCTCCAGAAAAGATGTCTATTTACTGATACCATCACCATGAAGGCGCAAGT 4138
    DB 1261 TGGCATCTCCAGAAAAGATGTCTATTTACTGATACCATCACCATGAAGGCGCAAGT 1320
    QY 4139 TTGAGTCCAGATTTAAACCAAG----- 4160
    DB 1321 TTGAGTCCAGATTTAAACCAAGCGAGTTCGATTTTAAAGAGAGCATGTGCGGGGAATCA 1380
    QY 4161 -----CAGGTTTCATGAAGGCTTAAAGTTATG 4186
    DB 1381 GGAGCTTAAAGACCTCAATAACAAACAACTTTTCAGGTTTCATGAAGGCTTAAAGTTATG 1440
    QY 4187 CATCATTAAGAAAGATTTATTTAGTCTCTTAAATCAAAATACCTCTGCAACGATCTT 4246
    DB 1441 CATCATTAAGAAAGATTTATTTAGTCTCTTAAATCAAAATACCTCTGCAACGATCTT 1500
  
```


TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2497)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saio, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashi, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers 1..2497 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:9630041G24" /db_xref="taxon:10090" /clone="9630041G24" /tissue types="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="16 days neonate" 24..>2495 /note="unnamed protein product; hypothetical ARM repeat structure containing protein (SCOP 48371, evidence: SCOP) putative" /codon_start=1 /protein_id="BAC29321.1" /db_xref="GI:26331182" /translation="MAAIGRSILKNLRIGRNDSCENVPLDLTREPSNLRILQNVAKLQVSNMRKGLHNNFYKLICDGHSEKLGPNVEDIICLRALLNEKEVRAA GLRALRYLIQDSSILQKLVKLDVYLARIQDQNEVERTQALRLVRMTITVNASL FFSVANSIANGVDQERDMVRACIAICELALQNEPVVALRGGLANTLKNVDC QLRINEALITATPSWAGIINLCPKNGSIQSLIGVLCIPNMEIRRGLELVYDI REARFLASKMGIITATPSWAGIINLCPKNGSIQSLIGVLCIPNMEIRRGLELVYDI FRPLPVWDEFLEALLSVDPGRFSDRLSDGFVAEAKTILPHRSPDLMDNYL ALILSAPFIRNGLEGLVEVITNSDDHISVRATILGELLHMANTILPHSHSHLCLP TLNMAASFDIPKEKRLAALNCLNRRMKRGPYPYSLLHLDHIQKAIATHHRK DQYLRVQKDFLVKDTDEALLINRDSQVLHKNLNDWNLGITLKNPNNLRNYK DEQLHRVRLLYFPKPSKYLALDIDLAKSLQTVGQCFTEFLLESDEQGYLE DLVKDI VQWLNASGVKPSRLQNGLLTTLQSLYFLFGLTSLCHPHGVKMLEKSVF QCLNLCLSNQDHLKLVSSLDYSPDGLARVLKSLTAA2DACLRYATKHLRVL RANVEFNWGIIELVTLQLDHKNKTISSSEALDILDNCEKXNHLALIOMKPALSHLG DKGLLLLRFLSPKPSYLNERYGV"
CDS	37..98 /note="unnamed protein product; hypothetical ARM repeat structure containing protein (SCOP 48371, evidence: SCOP) putative" /codon_start=1 /protein_id="BAC29321.1" /db_xref="GI:26331182" /translation="MAAIGRSILKNLRIGRNDSCENVPLDLTREPSNLRILQNVAKLQVSNMRKGLHNNFYKLICDGHSEKLGPNVEDIICLRALLNEKEVRAA GLRALRYLIQDSSILQKLVKLDVYLARIQDQNEVERTQALRLVRMTITVNASL FFSVANSIANGVDQERDMVRACIAICELALQNEPVVALRGGLANTLKNVDC QLRINEALITATPSWAGIINLCPKNGSIQSLIGVLCIPNMEIRRGLELVYDI REARFLASKMGIITATPSWAGIINLCPKNGSIQSLIGVLCIPNMEIRRGLELVYDI FRPLPVWDEFLEALLSVDPGRFSDRLSDGFVAEAKTILPHRSPDLMDNYL ALILSAPFIRNGLEGLVEVITNSDDHISVRATILGELLHMANTILPHSHSHLCLP TLNMAASFDIPKEKRLAALNCLNRRMKRGPYPYSLLHLDHIQKAIATHHRK DQYLRVQKDFLVKDTDEALLINRDSQVLHKNLNDWNLGITLKNPNNLRNYK DEQLHRVRLLYFPKPSKYLALDIDLAKSLQTVGQCFTEFLLESDEQGYLE DLVKDI VQWLNASGVKPSRLQNGLLTTLQSLYFLFGLTSLCHPHGVKMLEKSVF QCLNLCLSNQDHLKLVSSLDYSPDGLARVLKSLTAA2DACLRYATKHLRVL RANVEFNWGIIELVTLQLDHKNKTISSSEALDILDNCEKXNHLALIOMKPALSHLG DKGLLLLRFLSPKPSYLNERYGV"
ORIGIN	Query Match 37..98; Score 2146.6; DB 3; Length 2497; Best Local Similarity 91.2%; Pred. No. 0; Matches 2278; Conservative 0; Mismatches 219; Indels 0; Gaps 0; 5 GGTGTGACTGAACCGCTCAATATGCGCGCGATCGCGCGCGCTCTCTCTGAAGAACC 64 1 GGTGTGACTGAACCGCTCAATATGCGCGCGATCGCGCGCGCTCTCTCTGAAGAACC 60 65 TCCGAGTACGAGCGCGGAATGACAGCGCGCGAGGAGAACGTCCCGCTCGATCTGACCCGAG 124 61 TCCGAATACGAGCGCGGAATGACAGCGCGCGAGGAGAACGTCCCGCTCGATCTGACCCGAG 120 125 AACCTTCTGTATACTTAAGAGAGATTCTTCCAAAATGTGGCCAGATTGCGAGGAGTATCAA 184 121 AACCTTCTGTATACTTGGAGGAGATTCTTCCAAAATGTGGCCAAAATTCGAAGGAGTATCAA 180 185 ATATGAGAAGCTAGGCCATCTGAATACTTTACTTAAGCTTCTTTGTGATATTGGCCACA 244 181 ATATGAGAAGCTAGGCCATCTGAATACTTTACTTAAGCTTCTTTGGGATATTGGCCATA 240 245 GTGAAGAAAACTGGGCTTTCATCTATGAGGATATCAATAATTTGTTGCGGTAGCTTTAT 304 241 GTGAAGAAAACTGGGCTTTCATCTATGAGGATATCAATAATTTGTTGCGGTAGCTTTAT 300 305 TAAATGAAGCAAAAGTAGTCGAGCAGCAGCGGTACGAGCGCTTCGATATCTCATCAAG 364 301 TAAATGAGGCAAAAGTAGTAAGAGCAGCAGCGGTTCGAGCGCTTCGATATCTCATCAAG 360 365 ACTCCAGTATTTCCAGAGAGTCTTAAATTTGAAGTGGACTATTTAATAGCTAGGTGCA 424 361 ATTCCAGTATTTCCAGAGAGTCTTAAATTTGAAGTGGACTATTTAATAGCTAGGTGCA 420 425 TTGACATACAAAGCAGCAGCGGTAGAGAGGACACAGCACATTCGATAGTCAGAAGA 484 421 TTGACATTCAGCAGAGCAGCGGTGAGAGGACACAGCCCCCTTCGCTTAGTCAGAAGA 480 485 TGATTACTGTGAATGCTTCCTTCTTCTAGTTCTGTGACCAACTCAATTAATTTGCACTTG 544 481 TGATTACTGTGAATGCTTCCTTCTTCTAGTTCTGTGACCAACTCAATTAATTTGCGGTG 540 545 GAAATGATGAGCTTCAAGAAAGACAGAAATGTCGAGCATGATGATTCATCTCTGTG 604 541 GAAATGATGAGCTTCAAGAAAGACAGAAATGTCGAGCATGATGATTCATCTCTGTG 600 605 AACTAGCACCTTCAAGAAAGACAGAAATGTCGAGCATGATGATTCATCTCTGTG 664 601 AACTAGCACCTTCAAGAAAGACAGAAATGTCGAGCATGATGATTCATCTCTGTG 660 665 AAAATGTGATTTGATTCGCAATTAAGTCGAATAAATGAGGCCCTTAATTAATTTGTC 724 661 AAAATGTGATTTGATTCGCAATTAAGTCGAATAAATGAGGCCCTTAATTAATTTGTC 720 725 ACCTTCTTAATCATCAAGACAGCTCGGAGTATGTCGAGCTGATGATGATTTAGAGAGAA 784 721 ACCTTCTTAATCATCAAGACAGCTCGGAGTATGTCGAGCTGATGATGATTTAGAGAGAA 780 785 TTTTACGACCCCTATCTGATTTTTCATCTACAGACATAGTCCAGATACAGCTGAAGGACAG 844 781 TTTTACGACCCCTATCTGATTTTTCATCTACAGACATAGTCCAGATACAGCTGAAGGACAG 840 845 TCAAAGAAGACAGAGAAGCAGATTTCTAGCCAGTAAAAATGGGAATCATAGCAACATTC 904 841 TCAAAGAAGACAGAGAAGCAGATTTCTAGCCAGTAAAAATGGGAATCATAGCAACATTC 900 905 GATCATGGGAGGATTTAATTAATTTATGTAAACCTGGAATTTCTGGGATCCAGTCTCTAA 964 901 GCTCGTGGGAGGATTTAATTAATTTATGTAAGGCTCGGAATTTCTGGGATACAGTCTCTAA 960 965 TAGGAGTACTTTCATACCAATATGGAATAGGCGGAGTCTACTTGAAGTCTTTATG 1024 961 TTGGAGTACTTTCATACCAATATGGAATAGGCGGAGTCTGCTCGAAGTCTCTATG 1020 1025 ATATATTTCT 1084 1021 ACATCTTCT 1080 1085 TAGATCCAGGAGGTTTCCAAAGACAGTTGGAGGCTTTTCAGATGCTTTTGTGGCAGCTGAG 1144 1081 TAGATCCAGGAGGTTTCCAAAGACAGTTGGAGGCTTTTCCGATGCTTTTGTGGCAGCTGAG 1140

1145 CAAAACATTTCTCTCTCGTGCAGATCCAGGCGAGACCTCATGGAATAATTTTGG 1204
1141 CAAAACATTTCTCTCTCGTGCAGATCCAGGCGAGACCTCATGGAATAATTTGG 1200
1205 CACTGATCTCTCTCGATTTATTCGTAATGAGCTTTTATGAGGCTTCTAGTTGAGTGA 1264
1201 CTCTGATCTCTCTCGATTTATTCGTAATGAGCTTTTATGAGGCTTCTAGTTGAGTGA 1260
1265 CAAACAGTGTATGATATCTAGTTAGAGCTACCATCTTTTATGAGAGCTTTTACATA 1324
1261 CGAACAGTGTATGATATCTAGTTAGAGCTACCATCTTTTATGAGAGCTTTTACATA 1320
1325 TGGCAAAACAAATTTCTCTCTCATTTACATAGCCATCATTTTACACTCTCTTGGCCAAACCTTAA 1384
1321 TGGCAAAACAAATTTCTCTCTCATTTACATAGCCATCATTTTACACTCTCTTGGCCAAACCTTAA 1380
1385 TGAATATGCTGCATCTCTTGTATATCCCAAGGAAAGAGACTGCGAGGCGAGTGCAGCCT 1444
1381 TGAACATGCTGCCTCTCTTGTATATCCCAAGGAAAGAGGCTGCGAGGCGAGTGCAGCCT 1440
1445 TGAACATGCTGCCTCTCTTGTATATCCCAAGGAAAGAGGCTGCGAGGCGAGTGCAGCCT 1504
1441 TGAACATGCTGCCTCTCTTGTATATCCCAAGGAAAGAGGCTGCGAGGCGAGTGCAGCCT 1500
1505 ATTAGACACATATTTTCAAGAAAGCAATTTGCAACACAGCAAGGAGTCAAGTATCTCC 1564
1501 ATTAGATCACATTTTCAAGAAAGCAATTTGCAACACAGCAAGGAGTCAAGTATCTCC 1560
1565 GAGTTTCAAGAAAGATATATTTTCTTAAAGGATACAGAGGAGCTCTTTTAAATTAACCTTA 1624
1561 GAGTTTCAAGAAAGATATATTTTCTTAAAGGATACAGAGGAGCTCTTTTAAATTAACCTTA 1620
1625 GAGATGACCAAGCTCTTCAACATAAGAGAAATTTGAATGGAATTTGGAATCTTATAGGA 1684
1621 GAGACAGCAAGCTCTTCAACATAAGAGAAATTTGAATGGAATTTGGAATCTTATAGGA 1680
1685 CCATTTCTTAAGTGGCCAAATGTAATTTTAAAGAACTTAAAGATGAACAGTGTACACAGT 1744
1681 CCATCTTAAAGTGGCCAAATGTAATTTTAAAGAACTTAAAGATGAACAGTGTACACAGT 1740
1745 TTGTACGAAGACTTCTTTTATTTTCAAGCCAGCAGTAAATTTATATGTCACACCTGGATC 1804
1741 TTGTGCGCAGACTTCTTTTATTTTCAAGCCAGCAGCAAACTGTACGCTAGTCTGGATC 1800
1805 TGGATTTTCCAGGCAACAGCTCACCGTTGTAGTTGCGAGTTTACAGAAATTTCTTC 1864
1801 TGGACTTGGCCAAAGTCCAGACAGCTCACAGTTTGTGTGTTTACAGAAATTTCTGC 1860
1865 TTGAATCTCAAGAGATGGGCAAGGCTACTTGAAGATCTAGTAAAGGATATTTGTTCACT 1924
1861 TGAGTCTCAAGAGATGGGCAAGGATCTTGAAGATCTGTTGAAGATATTTGTTCACT 1920
1925 GGCTCAATGCTTCTATCTGGAATGAACCCGAAAGAGTCTTCAAAATATGTTTATGA 1984
1921 GGCTCAATGCTTCTATCTGGAATGAACCCGAAAGAGTCTTCAAAATATGTTTATGA 1980
1985 CCACCTTGTACACACTACTTTTATTTATTTGGAACACTTTCTTGGCCACCTCATGGAG 2044
1981 CCACCTTGTACACACTACTTTTATTTATTTGGAACACTTTCTTGTCTATCTCATGGAG 2040
2045 TTAATATGCTGAAATATCAGTGTATTTTCAAGTCTCTTCAATCTTCTCTCTTGA 2104
2041 TCAATATGCTGAAATATCAGTGTATTTTCAAGTCTCTTCAATCTTCTCTTGA 2100
2105 ACCAAGATCACTTGTGTAATTTTCTAGTTTCTAGCTTGGAGTATAGCAGATGGA 2164
2101 ACCAAGATCACTTGTGTAATTTTCTAGTTTCTAGCTTGGAGTATAGCAGATGGA 2160
2165 CTAGAGTCACTCTTCCAAATTTTAACTGCGAGTCTATGATGCGTGCAGCTCTATGCA 2224
2161 CCAGAGTCACTCTTCCAAATTTTAACTGCGAGTCTATGATGCGTGCAGCTCTATGCA 2220

2225 CAAAACATTTAAGGCTATTTATTTAGAGCTAATGTTGAATTTCTTAATATTTGGGAATG 2284
2221 CAAAACATTTAAGGCTATTTATTTAGAGCTAATGTTGAATTTCTTAACATTTGGGAATG 2280
2285 AGTTGTTAGTACCAGCTACATGATAAATAAACAAGATTTCTCTGAAAGCTCTTTGATA 2344
2281 AGTTACTAGTAACTCAGCTACATGATAAATAAACAAGATTTCTCTGAAAGCTCTGATA 2340
2345 TCCTCGATGAGCATGTGAGACAGGCGCAATCTTCATGCTCTCATTCAGATGAACAG 2404
2341 TTTCTTGAATGAGCTTGTGAAGACAGGCGCAATCTTCATGCTCTCATTCAGATGAACAG 2400
2405 CTTTATCCACCTTGGAGACAGGCTTGTCTCTCTGCTGAGATTTCTCTCATTTCCAA 2464
2401 CTTTATCCACCTTGGAGACAGGCTTGTCTCTCTGCTGAGATTTCTTCCATTTCCA 2460
2465 AAGGATTTTCTTATCTGAATGAAGAGGTTATGTAGC 2501
2461 AAGGATTTTCTTATCTGAATGAAGAGGTTATGTAGC 2497

RESULT 4
AK082586
LOCUS
DEFINITION
AK082586 3881 bp mRNA linear HTC 03-APR-2004
Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:C230066117 product:hypothetical ARM repeat
structure containing protein, full insert sequence.
AK082586
VERSION AK082586.1 GI:26100768
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE 4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
(bases 1 to 3881)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirokane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kohjiya, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

TITLE
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

FEATURES

source

1. 3881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:C230066117"
/db_xref="taxon:10090"
/clones="C230066117"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
24. 2581

misc_feature

/note="hypothetical ARM repeat structure containing protein (SCOP|48371, evidence: SCOP)

putative"

3868. 3873

polyA_signal

/note="putative"

polyA_site

3881

/note="putative"

ORIGIN

Query Match 37.8%; Score 2139.8; DB 3; Length 3881;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 2293; Conservative 0; Mismatches 222; Indels 10; Gaps 1;

Qy 5 GGTGTGACTGAACCCGCTCAATATGGCGCGATCGCGCGCGCTCTCTGAAGAACC 64
Db |||||
Qy 1 GGTGTGACTGAACCCGCTCAATATGGCGCGATCGCGCGCGCTCTCTGAAGAACC 60
Db |||||
Qy 65 TCCGAGTACGAGCGGCGAATGACAGCGCGGAGGAGACGTCCTCGATCTGACCCGAG 124
Db |||||
Qy 61 TCCGAAATACGAGGCGGGAATGACAGCGCGGAGGAGACGTCCTCGATCTGACCCGAG 120
Db |||||
Qy 125 AACCTTCTGATAACTTAAGAGAGATTCTCCAAAATGTGGCCAGATTGCGGGAGTATCAA 184
Db |||||
Qy 121 AACCTTCTGATAACTTAAGAGAGATTCTCCAAAATGTGGCCAGATTGCGGGAGTATCAA 180
Db |||||
Qy 185 ATATGAGAAAGCTAGGCCATCTGAATACTTTACTAAGCTTCTTTGTGATATTGGCCACA 244
Db |||||
Qy 181 ATATGAGAAAGCTAGGCCATCTGAATACTTTACTAAGCTTCTTTGGATATTGGCCATA 240
Db |||||
Qy 245 GTGAAGAAAACCTGGGCTTTCACATGAGGATATCAATAATTTGTTGGGGTACCTTTAT 304
Db |||||
Qy 241 GTGAAGAAAACCTGGGCTTTCATTAATGAGAGACATATAAATTTGTTGGGCTTAGCTTTAT 300
Db |||||

Qy 305 TAAATGAAGCAAAAGAAAGTGCAGCAGCAGCGGCTACGAGCGCTTCGATATCTCATCCCAAG 364
Db |||||
Qy 301 TAAATGAGGCAAAAGAAAGTAAAGAGCAGCGGGTTGCGAGCGCTGCGCTATCTCATCCCAAG 360
Db |||||
Qy 365 ACTTCAGTATTTCTCCAGAGGTGCTAAATTTGAAAGTGGACTATTTAATAGCTAGGTGCA 424
Db |||||
Qy 361 ATTCCAGTATTTCTCCGAAGGTGCTTAAACCTGAAAGTGGACTATTTAATAGCTAGGTGCA 420
Db |||||
Qy 425 TTGACATACAAACAGAGCAACGAGGTAGAGAGGACACAAAGCACTTCGATTAGCTCAGAAAAGA 484
Db |||||
Qy 421 TTGACATTCAGCAGAGCAACGAGGTGAGAGGACACAAAGCCCTTCGCTTAGTTCAGAAAAGA 480
Db |||||
Qy 485 TGATTACTGTGAATGCTTCCCTTGTCTAGTTCTGTGACCAACTCATTAATTTGCGAGTTG 544
Db |||||
Qy 481 TGATTACTGTGAATGCTTCCCTTGTCTAGTTCTGTGGCCAACTCATTAATTTGCGGTTG 540
Db |||||
Qy 545 GAAATGATGAGACTTCAAGAAAAGAGACAGATGCTCCGAGCATGCAATTTGCCATTTATCTGTG 604
Db |||||
Qy 541 GAAATGATGGGCTTTCAGGACGAGACAGAAATGGTCCGGGCTTGCATCGCCATCATCTGTG 600
Db |||||
Qy 605 AACTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAACACCATATTTGA 664
Db |||||
Qy 601 AACTAGCACTTCAGAACCCAGAGGTGGTAGCCCTTCGAGGAGGACTAAACACCATATCTGA 660
Db |||||
Qy 665 AAAATGTGATTTGATTTGCCAATTTAAGTCGAATTAATGAGGCCCTAATTTACTACAATTTTGC 724
Db |||||
Qy 661 AAAATGTGATTTGATTTGCCAATTTAAGTCGAATTAATGAGGCTCTCATTTACTACAATTTTGC 720
Db |||||
Qy 725 ACCTTCTTAATCATCCAAAGACTCGGCAGTATGTGCGAGCTGATGTAGAAATTTAGAGAA 784
Db |||||
Qy 721 ACCTTCTTAATCATCCAAAGACCCGCGAGTATGTGCGAGCTGATGTAGAAATTTAG 774
Db |||||
Qy 785 TTTTTCAGCACCTTACTACTGATTTTTCACCTACAGACATAGTCAGATACAGCTCAAGGACAGC 844
Db |||||
Qy 775 ----AGCACCTTACTACTGATTTTCCACTACAGACACAGTCAGATACGCTCGAGGACAGC 830
Db |||||
Qy 845 TCAAAGAGAGCAGAGAAAGCAGATTTTCTAGCCAGTAAATTTGGGAATCATATCAACATTTCC 904
Db |||||
Qy 831 TCAAAGAGAGCAGAGAAAGCAGATTTTCTAGCCAGTAAATTTGGGAATCATTTGCGACATTTCC 890
Db |||||
Qy 905 GATCATGGGAGGTATTTAATTTATTTATTTAACTGTAACCTGGAATTTCTGGGATCAGTCTCTAA 964
Db |||||
Qy 891 GCTCGTGGGAGGTATTTAATTTATTTATTTAAGCTGTAAGCTGGAATTTCTGGGATCAGTCTCTAA 950
Db |||||
Qy 965 TAGGAGTACTTTTCATACCAATATGGAATTAAGGCGAGGCTCTACTTTGAGTGTCTTTATG 1024
Db |||||
Qy 951 TTGGAGTACTTTGCATACCCGAATATGGAATTAAGGCGAGGCTCTGCTCGAAGTGTCTATG 1010
Db |||||
Qy 1025 ATATATTTCTCTTCTTCTACTCTGCTGAGTGTGAGGAGTTTCATAGAGACACTACTCAGTG 1084
Db |||||
Qy 1011 ACATCTTTCTCTCTCTCTCTGCGGTTGTGACGATGAGTTTCATAGAGACCTGCTCAGTG 1070
Db |||||
Qy 1085 TAGATCCAGGAGGTTTCCAGACAGTTTGGAGGCTTTTCAGATGGCTTTTGTGGCAGCTGAGG 1144
Db |||||
Qy 1071 TAGATCCAGGAGGATTTCCAGACAGTTTGGAGGCTTTCCGATGGCTTTTGTGGCAGCTGAGG 1130
Db |||||
Qy 1145 CAAAACATTTCTTCTCATCTGCGCAGATCCAGGCGAGACCTTCATCGATTAATTTATTTGG 1204
Db |||||
Qy 1131 CAAAACATTTCTTCTCTCTGCTGCTAGATCCAGGCGAGACCTTCATGGCAATTTATTTGG 1190
Db |||||
Qy 1205 CACTGATACTCTCTGCAATTTATTTTCGTAATGCACTTTTTCAGAGGCTTTCAGTGAAGTGATAA 1264
Db |||||
Qy 1191 CTCTGATACTCTCTGCAATTTATTTTCGTAATGCACTTTTTCAGAGGCTTTCAGTGAAGTGATAA 1250
Db |||||
Qy 1265 CAAAACAGTGAATCATPATCTCAGTTAGAGCTACCATCTCTTTTTCAGGAGAGCTTTTACATA 1324
Db |||||
Qy 1251 CGAACAGTGAACGACCATCTCAGTGAAGGCGACCATCTCTTTTAGGAGAGCTCTCGCATATA 1310
Db |||||
Qy 1325 TGGCAACACAAATTTCTTCTCATTTACATAGCCATCATTTTACTGCTTGGCCACCCCTAA 1384
Db |||||
Qy 1311 TGGCAACACCAATCTCTGCTCTCCTGATGACCCACCATCTACACTGCTTGGCCACCCCTCA 1370
Db |||||
Qy 1385 TGAATATGGTGCATCTCTTTGATATCCCAAGGAAAAGAGACTGCGGAGCCAGTGCAGCCT 1444
Db |||||

[illegible]

Db	2451	AAGATTTCCTTACCTGAATGAAGGGGTATGTTGCAAAACAGTTGGAAAAGTGGCACA	2511		
Qy	2525	GGGAA	2529		
Db	2511	AGGTA	2515		
RESULT 5	AY417033	1475 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	DEFINITION	Homo sapiens HCM6089 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.			
ACCESSION	AY417033	1 GI:39772993			
VERSION	AY417033.1	GI:39772993			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 1475)			
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Parriera, S., Wang, G., Zheng, X.H., White, T.J., Sminsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene tricos				
JOURNAL	Science	302	(5652),	1960-1963	(2003)
PUBMED	14671302				
REFERENCE	2	(bases 1 to 1475)			
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Parriera, S., Wang, G., Zheng, X.H., White, T.J., Sminsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003)	Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA			
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..1475	/organism="Homo sapiens"			
gene	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	<1..>1475				
ORIGIN	/locus_tag="HCM6089"				
Query Match	18.5%;	Score 1047.6;	DB 9;	Length 1475;	
Best Local Similarity	72.5%;	Pred. No. 5e-267;			
Matches 1066;	Conservative 0;	Mismatches 398;	Indels 6;	Gaps 1;	
Qy	125	AACCTTCGTGATCTTAAGAGAGATTCTCCAAATGTGCCAGATTGCGGGAGTATCAA	184		
Db	1	AACCTTCGTGATCTTAAGAGAGATTCTCCAAATGTGCCAGATTGCGGGAGTATCAA	60		
Qy	185	ATATGAGAAAGCTAGGCCATCTCAATACTTTACTAAGCTTCTTTGTGATATTGGCCACA	244		
Db	61	ATATGAGAAAGCTAGGCCATCTCAATACTTTACTAAGCTTCTTTGTGATATTGGCCACA	120		
Qy	245	GTGAAGAAAACCTGGGCTTTCACTATGAGATATCATAAATTCGTTTCGGTTAGCTTTAT	304		
Db	121	GTGAAGAAAACCTGGGCTTTCACTATGAGATATCATAAATTCGTTTCGGTTAGCTTTAT	180		
Qy	305	TAAATGAGCAAAAGAGTGCAGAGCAGCGGCTAGCGGCTTCGATATCTCATCCAAG	364		
Db	181	TAAATGAGCAAAAGAGTGCAGAGCAGCGGCTAGCGGCTTCGATATCTCATCCAAG	240		
Qy	365	ACTCCAGTATTTCCAGAGAGTGTCTAAATTTGAAAGTGGACTATTTAATAGCTAGTGCA	424		
Db	241	ACTCCAGTATTTCCAGAGAGTGTCTAAATTTGAAAGTGGACTATTTAATAGCTAGTGCA	300		
Qy	425	TTGACATACAAAGAGCAGGTTAGAGGACACAAGCATTTGATTTAGTCAGAAAGA	484		

Db 301 TTGACATACACAGAGCAACGAGGTAGAGAGGACACAAAGCACTTCGATTAGCTAGAAGA 360
Qy 485 TGATTACTGTGAATGCTTCTTGTCTTCTAGTCTTCTGTGACCAACTCAATTAATTCAGTTG 544
Db 361 TGAATTACTGTGAATGCTTCTTGTCTTCTAGTCTTCTGTGACCAACTCAATTAATTCAGTTG 420
Qy 545 GAAATGATGAGACTTCAAGAAAGACACGAATGGTCCGAGCATGCAATTCGCCATTAATTCGTG 604
Db 421 GAAATGATGAGACTTCAAGAAAGACACGAATGGTCCGAGCATGCAATTCGCCATTAATTCGTG 480
Qy 605 AACTAGCACTTTCAGAACTCCAGAGTGGTGGCCCTTCGAGGAGGAGTAAACACCATATTTGA 664
Db 481 NNN 540
Qy 665 AAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
Db 541 AAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 725 ACCTTCTTAATCATCCAAAGACTCGGCGATGATGCGAGCTGATGATGATGATGATGATGATGAT 784
Db 601 ACCTTCTTAATCATCCAAAGACTCGGCGATGATGCGAGCTGATGATGATGATGATGATGATGAT 660
Qy 785 TTTTAGCACCTATACCTGATTTTCACTACACATAGTCCAGATACACAGCTGAAGCAGC 844
Db 661 NNN 720
Qy 845 TCAAAGAAGACAGAGAAGCAGATTTCTAGCCAGTAAATGCGAATCATAGCAACATCC 904
Db 721 NNN 780
Qy 905 GATCATGGCGAGGTATTATTAATTTATGTAACCTGGAAATTTCTGGGATCCAGTCTCTAA 964
Db 781 NNNNNNNNNNNNGTATTATTAATTTATGTAACCTGGAAATTTCTGGGATCCAGTCTCTAA 840
Qy 965 TAGGAGTACTTTCATACCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 1024
Db 841 TAGGAGTACTTTCATACCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 894
Qy 1025 ATATATTTCGTCTCTCTACCTGCTGACTGAGGAGTTTCATAGAGACACTACTCAGTG 1084
Db 895 NNN 954
Qy 1085 TAGATCCAGGAGGTTTCAAGACAGTGTGGAGCTTTTTCAGATGGCTTTTGTGGCAGCTGAG 1144
Db 955 NNNNNNNNNNGAGGTTTCAAGACAGTGTGGAGCTTTTTCAGATGGCTTTTGTGGCAGCTGAG 1014
Qy 1145 CAAAACTATTCTTCTCATCGTCCAGATCCAGCCAGCTCATGCGATAATTTATG 1204
Db 1015 CAAAACTATTCTTCTCATCGTCCAGATCCAGCCAGCTCATGCGATAATTTATG 1074
Qy 1205 CACTGATACTCTCTGCTATTTCGTAATGCACTTTTAGAGGCTTAGTTGAACTGATAA 1264
Db 1075 NNN 1134
Qy 1265 CAAACAGTGATGATCATATCTCAGTTAGAGCTACCATCTTTTAGGAGAGCTTTTACATA 1324
Db 1135 NNN 1194
Qy 1325 TGGCAACACAAATCTTCTCATTCATACATAGCCATCATTTACACTGCTTCCCAACCTTAA 1384
Db 1195 NNGCAACACAAATCTTCTCATTCATACATAGCCATCATTTACACTGCTTCCCAACCTTAA 1254
Qy 1385 TGAATATGGCTGCTATCTTTGATATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCCT 1444
Db 1255 TGAATATGGCTGCTATCTTTGATATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCCT 1314
Qy 1445 TGAATCTTTTAAACGCTTCCATGAAATGAAAGAAACGAGGACCTTAAGCCTTATAGTCTTC 1504
Db 1315 TGAATCTTTTAAACGCTTCCATGAAATGAAAGAAACGAGGACCTTAAGCCTTATAGTCTTC 1374
Qy 1505 ATTTAGACCACTATTTCAGAAAGCAATTTGCAACACACACAGAAACGCGGATCAGTATCTCC 1564
Db 1375 ATTTAGACCACTATTTCAGAAAGCAATTTGCAACACACACAGAAACGCGGATCAGTATCTCC 1434

Qy 1565 GAGTTCAGAAAGATATATTTATCTCTTAAGG 1594
Db 1435 GAGTTCAGAAAGATATATTTATCTCTTAAGG 1464

RESULT 6
AK075662
LOCUS
DEFINITION
Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
enriched library, clone:1110055K10 product:unknown EST, full insert
sequence.
AK075662 1977 bp mRNA
AK075662.1 GI:26344462
VERSION
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1977)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.

FEATURES	SOURCE
-----------------	---------------

Source

1. 1977

```

1. 1977
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:110055K10"
/db_xref="taxon:10090"
/clone="110055K10"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
<1. 1089
/notes="unnamed protein product; putative
unknown EST (GB|AV014646, evidence: BLASTN, 99%,
match=698)"

```

ORIGIN

Query Match	17.78;	Score 1001.6;	DB 3;	Length 1977;
Best Local Similarity	82.2%;	Pred. No. 9.9e-255;		
Matches 1290;	Conservative 0;	Mismatches 254;	Indels 26;	Gaps 11;
QY	4069	TCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTACTGATACCATCACCATGAAG	4128	
DB	1	TCGGAAGCTTTGGCGTCCGCGCAAAAGATGTGCTTTTACCGACACCATCACCATGAAG	60	
QY	4129	GCCAAAGTTTTGAGTCCAGATTAAACACCAAGCAGGTTTCATGAAGGCTTTAAGTTATGCA	4188	
DB	61	GCCAAACAGCTTTGAATCCAGACTGACGCCAAGCAGGTTTATGAAGGCTTTAAGTTATGCG	120	
QY	4189	TCATTAGATAAAGAAGATTTTATTGAGTCTTATTAAATCAAATAACCCCTGCAACGATCTTCC	4248	
DB	121	TCATTAGATAAAGAGATTTTACTAAGTCTTATTAAATCAACATACCCCTCCAGCGCTCTCC	180	
QY	4249	TCAGTGGCGTCCATGGTGTCCAGTGCACATATGCGGGGTTTCAGATGATTACATTTGTCCTT	4308	
DB	181	TCAGTTAGGTCTATGGTGTCCAGTGTCTACATATGGAGGTTTCCGATGACTACATTTGCACTT	240	
QY	4309	GCCTCCCGGTGGATATAAATGATATATCCAGGTAAGGATATTCCTCTATTTTCAGACA	4368	
DB	241	GCCTCTCCAGTAGACATCAATGATCTTTCAGATAAAGGATGTTCTTATTTTCATCA	300	
QY	4369	AAAAACATACCAACCATGATGATTCGAGGTGCAGAGCATTTGCCCATGATCAGGAGGT	4428	
DB	301	AAGCACGTGCTCCACCAGATGACCGAGGTGCAAGGATGTTTTCATGATGAGCAGGT	360	
QY	4429	CTTCCATCTCGAACTGGAGTCTGTGTAATAAATCTCTTTTCACTTGTCTACGACGCGAGTG	4488	
DB	361	CTTTTCATCTGGAGCCGGTGTCTCGTAAGAACTCTTTTCACTTGTCTGCGGACGAGTG	420	
QY	4489	AGTCTTACGGAAATAATGAATTCATCCATTCAGATGCCCTCTCTGTTTTTTAGAAAGTACA	4548	

LOCUS BG623200 856 bp mRNA linear EST 18-APR-2001
DEFINITION 602648109F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4769547 5',
mRNA sequence.
ACCESSION BG623200
VERSION BG623200.1 GI:13674571
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1635 row: a column: 04
High quality sequence stop: 680.
FEATURES
source
1..856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4769547"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 79"
/note="Organ: Placenta; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggccattatggcc); Site 2: SfiI
(ggccattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATCTGAGCGCGGCGGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
Kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

Query Match 13.3%; Score 755.2; DB 4; Length 856;
Best Local Similarity 96.7%; Pred. No. 3.3e-189;
Matches 824; Conservative 0; Mismatches 23; Indels 5; Gaps 5;
QY 5 GGTGTGACTGAACCCGGTCAATATGCGCGGCGATCGCGCGCGCTCTCTGAAGAACC 64
DB 1 GGTGTGACTGAACCCGGTCAATATGCGCGGCGATCGCGCGCGCTCTCTGAAGAACC 60
QY 65 TCCGAGTACGAGGCGGGAATGACAGCGGCGAGGAGAACGTCCCGTGGATCTGACCCCGAG 124
DB 61 TCCGAGTACGAGGCGGGAATGACAGCGGCGAGGAGAACGTCCCGTGGATCTGACCCCGAG 120
QY 125 AACCTTCTGATTAACCTTAAGAGAGATCTCCAAAATGTGGCCAGATGCGAGGAGATCA 184
DB 121 AACCTTCTGATTAACCTTAAGAGAGATCTCCAAAATGTGGCCAGATGCGAGGAGATCA 180
QY 185 ATATGAGAAAGCTAGGCCATCTGAATACTTACTTAAGCTTCTTTGTGATATGGCCACA 244
DB 181 ATATGAGAAAGCTAGGCCATCTGAATACTTACTTAAGCTTCTTTGTGATATGGCCACA 240
QY 245 GTGAAGAAAGCTAGGCCATCTGAATACTTACTTAAGCTTCTTTGTGATATGGCCACA 304
DB 241 GTGAAGAAAGCTAGGCCATCTGAATACTTACTTAAGCTTCTTTGTGATATGGCCACA 300
QY 305 TAAATGAACAAAGAGTGGCAGCAGCAGGCGGTACGAGCGTTCGATATCTCATCCCAAG 364
DB 301 TAAATGAACAAAGAGTGGCAGCAGCAGGCGGTACGAGCGTTCGATATCTCATCCCAAG 360

QY 365 ACTCCAGTATTCTCAGAAGGTGCTAAATTTGAAAGTGGACTATTATATAGTAGGTGCA 424
DB 361 ACTCCAGTATTCTCAGAAGGTGCTAAATTTGAAAGTGGACTATTATATAGTAGGTGCA 420
QY 425 TTGACATACAAACAGAGCAACGAGGTAGAGAGACACAAAGCACTTCGATTGATCGAAGA 484
DB 421 TTGACATACAAACAGAGCAACGAGGTAGAGAGACACAAAGCACTTCGATTGATCGAAGA 480
QY 485 TGATTACTGTGAATGCTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 544
DB 481 TGATTACTGTGAATGCTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 545 GAAATGATGGACTCTCAAGAAAGACAGAAATGCTCGAGACATGATTCGCCATTAATCTCTG 604
DB 541 GAAATGATGGACTCTCAAGAAAGACAGAAATGCTCGAGACATGATTCGCCATTAATCTCTG 600
QY 605 AACTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACT-AAACACCATATG 663
DB 601 AACTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACT-AAACACCATATG 660
QY 664 AA-AAATGTGATGATTCGCAATTAAGTCGAAT-AAATGAGGCCCTTAATTAACAATTT 721
DB 661 AA-AAATGTGATGATTCGCAATTAAGTCGAAT-AAATGAGGCCCTTAATTAACAATTT 720
QY 722 TGCACCTTCTTAATCATCCAAAGACTCGGCAGATG-TGCAGAGCTGATGTAGAATTAG 780
DB 721 TGCACCTTCTTAATCATCCAAAGACTCGGCAGATG-TGCAGAGCTGATGTAGAATTAG 780
QY 781 AGAATTTAGCAACCTTACTGATTTTCACTACAGACATAGTCAGATACAGCTGAAGA 840
DB 781 AGAATTTAGCAACCTTACTGATTTTCACTACAGACATAGTCAGATACAGCTGAAGA 839
QY 841 CAGCTCAAGAA 852
DB 840 CAGTTCAAGAAA 851

RESULT 11
CD653846 878 bp mRNA linear EST 18-JUN-2003
LOCUS AGSNCOURT 14556996 NIA Human HI Embryonic Stem Cell cDNA Library
DEFINITION (Long) Homo sapiens cDNA clone IMAGE:30425290 5', mRNA sequence.
ACCESSION CD653846
VERSION CD653846.1 GI:31892651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 878)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM509 row: e column: 11
High quality sequence stop: 704.
FEATURES
source
1..878
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:30425290"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on ME9 feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSA-4, OCT3, OCT4, REX1, UFG, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGGCCCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker Lc-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricion-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricion-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 13.3%; Score 753.2; DB 6; Length 878;
 Best Local Similarity 94.9%; Pred. No. 1.1e-188;
 Matches 782; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

Qy 1188 CATGGATAATTTATTTGGCACTGATACCTCTCGCATTTATTCGTATGACCTTTTAGAGGG 1247
 Db 11 CATGGATAATTTATTTGGCACTGATACCTCTCGCATTTATTCGTATGACCTTTTAGAGGG 70

Qy 1248 TCTAGTTGAAGTATAACAAACAGTGATGATCATATCTCAGTTAGACTACCATCTCTTTT 1307
 Db 71 TCTAGTTGAAGTATAACAAACAGTGATGATCATATCTCAGTTAGACTACCATCTCTTTT 130

Qy 1308 AGGAGAGCTTTTACATATGCGAACAACATCTTCTTCATTCATTCATAGCCATCATTTACA 1367
 Db 131 AGGAGAGCTTTTACATATGCGAACAACATCTTCTTCATTCATTCATAGCCATCATTTACA 190

Qy 1368 CTGCTTGCACACCTTAATGATATGCTGTCATCTTTGATATCCCAAGGAAAGAGACT 1427
 Db 191 CTGCTTGCACACCTTAATGATATGCTGTCATCTTTGATATCCCAAGGAAAGAGACT 250

Qy 1428 GCGAGCCAGTGCAGCTTTGAACTGTTTAAACGCTTCCATGAATGAAGAACGAGGACC 1487
 Db 251 GCGAGCCAGTGCAGCTTTGAACTGTTTAAACGCTTCCATGAATGAAGAACGAGGACC 310

Qy 1488 TAAGCCTTATAGTCTTCATTTAGACCATATTTTTCAGAAAGCAATTCAGAACACACAGAA 1547
 Db 311 TAAGCCTTATAGTCTTCATTTAGACCATATTTTTCAGAAAGCAATTCAGAACACACAGAA 370

Qy 1548 ACGGATCAGTATCTCCGAGTTTCAAGAGATATATTTTCTTAAAGTACAGAGGAGC 1607
 Db 371 ACGGATCAGTATCTCCGAGTTTCAAGAGATATATTTTCTTAAAGTACAGAGGAGC 430

Qy 1608 TCTTTTAATTAACCTTAGAGATAGCAAGTCTCTTCAACATAAAGAGATCTTGAATGGAA 1667

Db 431 TCTTTTAATTAACCTTAGAGATAGCAAGTCTCTTCAACATAAAGAGATCTTGAATGGAA 490
 Qy 1668 TTGGAATCTTATAGGAGCACTTTCTTAAGTGCCCAATTAATCTTAAGAACTATAAAGA 1727
 Db 491 TTGGAATCTTATAGGAGCACTTTCTTAAGTGCCCAATTAATCTTAAGAACTATAAAGA 550
 Qy 1728 TGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTTACAAGCCAGCAGTAAAT 1787
 Db 551 TGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTTACAAGCCAGCAGTAAAT 610
 Qy 1788 ATATGCCAACCTGATCTGGATTTTGCAGGCGCAACAGCTCAGGTTGTAGGTGGCA 1847
 Db 611 ATATGCCAACCTGATCTGGATTTTGCAGGCGCAACAGCTCAGGTTGTAGGTGGCA 670
 Qy 1848 GTTTACAGAAATTTCTTTGAATCTGAAGAGATGGGCAAGCTACTTTAGAGATCTAGT 1907
 Db 671 GTTTACAGAAATTTCTTTGAATCTGAAGAGATGGGCAAGCTACTTTAGAGATCTAGT 730
 Qy 1908 AAAGGATATTTGTCAGTGGCTCAATGCTTCATCTGGAATGAACCCGAAAGAGTCTTCA 1967
 Db 731 NAAGGATATTTGTCAGTGGCTCATGCTCATCTGGAATGAACCCGAAAGAGTCT--C 788
 Qy 1968 AAATAAGTGTATTATGACCCCTTGTAGTCAACACTACTTTTAT 2011
 Db 789 ANNATATGTTNNATGACACCTAGTCACTACTTTTATTTAT 832

RESULT 12
 BI601569
 LOCUS 603250782F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5302244 5',
 DEFINITION mRNA sequence.
 ACCESSION BI601569
 VERSION BI601569.1 GI:15494508
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 831)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: csapbs-x@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1765 row: d column: 21
 High quality sequence stop: 771.

FEATURES
source

1..831
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5302244"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_96"
 notes="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is

QY 5052 ATTATTTCAAGATGTACAGTTTCTACAAATGCATGAAGACAGAGGCTGTGTTGGCAAC 5111
 DB 541 ATTATTTCAAGATGTACAGTTTCTACAAATGCATGAAGACAGAGGCTGTGTTAGCAAC 600
 QY 5112 ACCACCAAGCAACTATAGTTGATACATCTGCTGATCCTGACCTCATATTTATGATGG 5171
 DB 601 ACCACCAAGCAACTATAGTTGATACATCTGCTGATCCTGACCTCATATTTATGATGG 660
 QY 5172 ATATAGATACATACATATATATATTTCAATTTGTTGATTTCTTAAAGCCTCAGAAATAC 5231
 DB 661 ATATAGATATATAGTATATATATTTCAATTTGTTGATTTCTTAAAGCCTCAG-AAATAC 719
 QY 5232 GACTGACTAGGACAGCAAGACAGAGTA 5259
 DB 720 GACTGACTAGGACAGCAAGACAGAGTA 747

RESULT 14
 BG414183/c
 LOCUS
 DEFINITION
 GKP2-237 GKP2 subtraction cDNA library Mus musculus linear EST 13-MAR-2003
 GKP2-237 5' and 3' similar to cytosolic regulator in G protein
 mediated adenylyl cyclase activation, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 867)
 Identification of thirty-six genes whose expression is markedly
 affected by beta cell differentiation and de-differentiation
 Unpublished (2001)
 Contact: Magnuson, M.A.
 Department of Molecular Physiology and Biophysics
 Vanderbilt University
 747 LH, 21st Ave at Garland Ave. Nashville, TN 37232-0615, USA
 Fax: 615 322 7236
 Email: mark.magnuson@mcmail.vanderbilt.edu
 PCR Primers
 FORWARD: M13 forward
 BACKWARD: M13 reverse
 Insert Length: 867 Std Error: 0.00
 Seq primer: T7, M13 reverse, custom
 High quality sequence stop: 867
 POLYA=No.

FEATURES
 source
 1. .867
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J x DBA/2"
 /db_xref="taxon:10090"
 /clone="GKP2-237"
 /cell_type="Pancreatic endocrine tumor"
 /cell_line="GKP2"
 /lab_host="TOP10 (Invitrogen, Carlsbad, CA)"
 /clone_lib="GKP2 subtraction cDNA library"
 /notes="Organ: Pancreas; Vector: pCR 2.1-TOPO (Invitrogen,
 Carlsbad, CA); Site 1: EcoRI; Site 2: BamHI; mRNA from
 GKP2 and GKP4 cells was used to perform subtraction
 suppression hybridization. Genes that are more highly
 expressed in GKP2 cells (an insulinoma derived cell,
 compared to GKP4 cells (a precursor cell line) were
 cloned pCR 2.1-TOPO and transformed into TOP10 competent
 bacteria."

ORIGIN

Query Match 12.5%; Score 710.2; DB 4; Length 867;
 Best Local Similarity 88.7%; Pred. No. 3.2e-177;
 Matches 769; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1790 ATGCAACCTGGATCTGGATTTTCCCAAGGCCAAACAGCTCACGGTTGTAGTTGCCAGT 1849
 DB 867 ACGTAGTCTGGATCTGGACTTGGCCAAAGTCCAGAGCAGCTCACAGTTGTGTGTCACT 808
 QY 1850 TTACAGAAATTTCTTCTGAAATCTGAAGAGGATGGCAAGGCTACTTAAAGACTTAGTAA 1909
 DB 807 TTACAGAAATTTCTGCTCGAGTCTGAAGAGGATGGCAAGGATACTTAAAGACTCTCGTGA 748
 QY 1910 AGGATATTTGTTCTGAGTGCTCAATGCTTCACTGGAATGAACCCGGAAGAGTCTTCAAA 1969
 DB 747 AAGATATTTGTTCTGAGTGCTCAATGCTTCACTGGAATGAACCCGGAAGAGTCTTCAAA 688
 QY 1970 ATAATGTTTATTTAGCACCACTCTTAGTCAACACTACTTTTATTTATTTATTTGGAACACTTCTT 2029
 DB 687 ATAATGCTTACTTGACCACTCTTAGTCAACACTACTTTTATTTATTTATTTGGAACACTTCTT 628
 QY 2030 GCCACCTCATGGAGTTAAATATGCTGGAATAATGCTGGAATAATGCTGAGTGTCTCTTAAATC 2089
 DB 627 GTCATCTCATGGAGTCAAAATGCTGGAATAATGCTGAGTGTCTCTTAAATC 568
 QY 2090 TTTGCTCTTGAAGAAACCAAGATCAGTCTGCTTAAACCTTACTGTTCTTAGCTTGACATATA 2149
 DB 567 TTTGTTCTTGAAGAAACCAAGATCAGTCTGATAAACTCAGTGTCTTCCAGCTGACATATA 508
 QY 2150 GCAGAGATGGAATGGCTAGAGTCACTCTTCCAAAAATTTTAACTGACAGTACTGATGCTT 2209
 DB 507 GCAGAGATGGAATGAGCCAGAGTCACTCTTCCAAAAATTTTAACTGACAGTACTGATGCTT 448
 QY 2210 GCAGACTTATGCAACAAAAACATTTAAGGATATTTATGAGAGTAAATGTTGAATCTTTA 2269
 DB 447 GCAGACTTATGCAACAAAAACATTTAAGGATTTTATGAGAGTAAATGTTGAATCTTTA 388
 QY 2270 ATAAATGGGAAATGAGTTCTTACTGAGCCAGCTTACATGATATAAAACAAAACGATTTTCT 2329
 DB 387 ACAATTTGGGAAATGAGTTCTTACTGAGTAACTCAGCTACATGATAAAACAAAACAAATTTCTT 328
 QY 2330 CTGAAGCTCTTGATATCTCGATGAAGCATGTGAAGACAAAGGCCAAATCTTCATGCTCTCA 2389
 DB 327 CTGAAGCTCTTGATATCTCGATGAAGCATGTGAAGACAAAGGCCAAATCTTCATGCTCTCA 268
 QY 2390 TTCAAGTAAACCCAGCGTTATCCACCTTGGAGACAAAGGTTGCTCTCTGCTGAGAT 2449
 DB 267 TTCAAGTAAACCCAGCGTTATCCACCTTGGAGACAAAGGTTGCTCTCTGCTGAGAT 208
 QY 2450 TTCTCTCCATTTCCAAAGGATTTCTTCTGATGAAGAGGTTATGTTAGCAAAACAAT 2509
 DB 207 TTCTTTCCATTTCCGAAAGGATTTTCTTACCTGATGAAGAGGTTATGTTAGCAAAACAAT 148
 QY 2510 TGGAAAGTGGCACAGGGAATCAACCTCCAAATATGTTGACTTGAATGAGGAACAACCTCA 2569
 DB 147 TGGAAAGTGGCACAGGGAATCAACCTCCAAATATGTTGACTTGAATGAGGAACAACCTCA 88
 QY 2570 ATGAAGCATTACTACTTACCGGAAGCTGTTGATGTTGATGATAAATATGTTCTGCTGGAGTA 2629
 DB 87 ATGAAGCCTTCACTACATACAGGAAGCTTATGATGCGGATAAATACGTTTCTGCTGAGTA 28
 QY 2630 ACCAAGATTTACAGGCTGCTCAGTCT 2656
 DB 27 ATCAAGACTACAGGCACTCTATGTTCT 1

RESULT 15

AK029196

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK029196 4697 bp mRNA linear HTC 03-APR-2004
 Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
 library, clone:483143G18 product:hypothetical ARM repeat structure
 containing protein, full insert sequence.

AK029196.1 GI:26325179

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Qy	725	ACCTTCTTAATCATCAAAAGACTCGGAGTATGTGCGAGTGATGTTAGAATTAGAG	780
Dδ	722	ACCTTCTTAATCATCAAAAGACCCGGAGTATGTGCGAGTGATGTTAGAATTAGAG	777

Search completed: April 9, 2005, 13:24:16
Job time : 11596.7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 17:46:42 ; Search time 2039.44 Seconds
(without alignments)
16824.125 Million cell updates/sec

Title: US-10-782-244-1

Perfect score: 5662

Sequence: 1 agcgggtgtgactgaacc.....gctggtattagagagggaac 5662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5662	100.0	5662	18	US-10-782-244-1
2	5127	90.6	5127	18	US-10-782-244-2
3	2244.2	39.6	4657	10	US-09-814-353-19068
4	2239.4	39.6	4601	18	US-10-357-930-20181
5	2239.4	39.6	4602	18	US-10-357-930-26008
6	465.2	8.2	570	10	US-09-814-353-11
7	465.2	8.2	570	10	US-09-814-353-6409
8	385.6	6.8	452	18	US-10-357-930-9378
9	383.4	6.8	427	18	US-10-357-930-30581
10	383.4	6.8	427	18	US-10-357-930-39551
11	336.4	5.9	416	18	US-10-357-930-209

12	334	5.9	341	10	US-09-814-353-12794	Sequence 12794, A
13	321	5.7	450	9	US-09-933-797-314	Sequence 714, Appl
14	280.2	4.9	726	9	US-09-910-943-73	Sequence 73, Appl
15	186	3.3	492	9	US-09-783-590-8386	Sequence 8386, Ap
16	166.8	2.9	567	16	US-10-029-386-8811	Sequence 8811, Ap
17	163.6	2.9	170	16	US-10-029-386-22537	Sequence 22537, A
18	142.2	2.5	1698	13	US-10-027-632-257900	Sequence 257900,
19	142.2	2.5	1698	13	US-10-027-632-257900	Sequence 257900,
20	88.6	1.6	405	17	US-10-242-535A-44306	Sequence 44306, A
21	88.6	1.6	405	17	US-10-242-535A-44306	Sequence 44306, A
22	77.4	1.4	289	10	US-09-814-353-18070	Sequence 18070, A
23	73	1.3	185	10	US-09-814-353-21600	Sequence 21600, A
24	65.2	1.2	170	17	US-10-242-535A-22309	Sequence 22309, A
25	65.2	1.2	170	17	US-10-242-535A-22309	Sequence 22309, A
26	60.8	1.1	280	18	US-10-085-783A-8321	Sequence 8321, Ap
27	57	1.0	185	10	US-09-814-353-5399	Sequence 5399, Ap
28	57	1.0	185	10	US-09-814-353-11886	Sequence 11886, A
29	53.2	0.9	551	18	US-10-363-345A-7393	Sequence 7393, Ap
30	53.2	0.9	551	18	US-10-363-345A-7393	Sequence 7394, Ap
31	53.2	0.9	551	19	US-10-363-483A-7393	Sequence 7393, Ap
32	53.2	0.9	551	19	US-10-363-483A-7393	Sequence 7394, Ap
33	51.4	0.9	551	18	US-10-363-345A-7395	Sequence 7395, Ap
34	51.4	0.9	551	18	US-10-363-345A-7396	Sequence 7396, Ap
35	51.4	0.9	551	19	US-10-363-483A-7395	Sequence 7395, Ap
36	51.4	0.9	551	19	US-10-363-483A-7396	Sequence 7396, Ap
37	51.2	0.9	530	18	US-10-437-963-76169	Sequence 76169, A
38	45.8	0.8	226215	13	US-10-087-192-1948	Sequence 1948, Ap
39	44.8	0.8	1265	13	US-10-027-632-261722	Sequence 261722,
40	44.8	0.8	1265	17	US-10-027-632-261722	Sequence 261722,
41	41.8	0.7	37515	18	US-10-433-793-28	Sequence 28, Appl
42	41	0.7	454	10	US-09-918-995-15405	Sequence 15405, A
43	41	0.7	1359	13	US-10-027-632-166149	Sequence 166149,
44	41	0.7	1359	17	US-10-027-632-166149	Sequence 166149,
45	41	0.7	136726	17	US-10-085-117-244	Sequence 244, App

ALIGNMENTS

RESULT 1

US-10-782-244-1

; Sequence 1, Application US/10782244

; Publication No. US20040253677A1

; GENERAL INFORMATION:

; APPLICANT: Sabatini, David M.

; APPLICANT: Kim, Do-Hyung

; APPLICANT: Sarbassov, Dos D.

; TITLE OF INVENTION: mTOR KINASE-ASSOCIATED PROTEINS

; FILE REFERENCE: WBL-P01-010

; CURRENT APPLICATION NUMBER: US/10/782,244

; CURRENT FILING DATE: 2004-02-18

; PRIOR APPLICATION NUMBER: US 60/448,035

; PRIOR FILING DATE: 2003-02-18

; NUMBER OF SEQ ID NOS: 51

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 5662

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-782-244-1

Query Match	100.0%	Score 5662;	DB 18;	Length 5662;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 5662;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	AGCGGGTGTGACTGAACCCGTCATATGCGCGGCGATCGCGCGCGCTCTCTGAAG	60	
Db	1	AGCGGGTGTGACTGAACCCGTCATATGCGCGGCGATCGCGCGCGCTCTCTGAAG	60	
QY	61	ACCTTCGAGTACGAGCGGGAATGACAGCGCGGAGAGAACCTCCGCTGGATCTGACC	120	
Db	61	AACTTCGAGTACGAGCGGGAATGACAGCGCGGAGAGAACCTCCGCTGGATCTGACC	120	

Db 2281 ATTGAGTTGTTAGTGACCCAGCTACATGATATAAACAACAAACGATTTCTCTGAGCTCTT 2340
Qy 2341 GATATCTCGATGAAGCATGTGAAGCAAGGCCAATCTTTCATGTCTCTCATTCAGATGAAA 2400
Db 2341 GATATCTCGATGAAGCATGTGAAGCAAGGCCAATCTTTCATGTCTCTCATTCAGATGAAA 2400
Qy 2401 CCAGGTTATCCCACTTGGAGACAAGGGTTGCTTCTCCTGCTGAGATTTCTCTCCATTT 2460
Db 2401 CCAGGTTATCCCACTTGGAGACAAGGGTTGCTTCTCCTGCTGAGATTTCTCTCCATTT 2460
Qy 2461 CCAAAAGGATTTTCTCTATCTGAATGAAGAGGTTATGTAGCAAAACAATTCGAAAAGTGG 2520
Db 2461 CCAAAAGGATTTTCTCTATCTGAATGAAGAGGTTATGTAGCAAAACAATTCGAAAAGTGG 2520
Qy 2521 CACAGGGGAATACAACTCCAAATATGTTGCACTTGATTTGAGGAACAACCTCAATGAAGCACTT 2580
Db 2521 CACAGGGGAATACAACTCCAAATATGTTGCACTTGATTTGAGGAACAACCTCAATGAAGCACTT 2580
Qy 2581 ACTACTTACCGGAAGCCTGTTGATGGTGATAACTATGTTCTCGAGGTAAACCAAGATTA 2640
Db 2581 ACTACTTACCGGAAGCCTGTTGATGGTGATAACTATGTTCTCGAGGTAAACCAAGATTA 2640
Qy 2641 CAGCGTCTCAGCTCTACTCGCTTATACACCTTTTATGGACAACCTAGTACACCAATAAACA 2700
Db 2641 CAGCGTCTCAGCTCTACTCGCTTATACACCTTTTATGGACAACCTAGTACACCAATAAACA 2700
Qy 2701 GGCTGCCATTTGTTGGAGTACAGAATATTTATCAGAACTCTGTCTGTAATGTTGTTGTA 2760
Db 2701 GGCTGCCATTTGTTGGAGTACAGAATATTTATCAGAACTCTGTCTGTAATGTTGTTGTA 2760
Qy 2761 CCAGATTTGATTAAGTGGGAAGAAATTAATAAACTGAAAGCATCTTTTGGGCGCTTGGGA 2820
Db 2761 CCAGATTTGATTAAGTGGGAAGAAATTAATAAACTGAAAGCATCTTTTGGGCGCTTGGGA 2820
Qy 2821 AATATCGCTCATCAAAATTTGGGCTCAATTTGCTTACAGGAAGAAACGTTGATTTCCAGAT 2880
Db 2821 AATATCGCTCATCAAAATTTGGGCTCAATTTGCTTACAGGAAGAAACGTTGATTTCCAGAT 2880
Qy 2881 ATACTAAAACTTGCAGAAACAGTGTGAAGTTCTTTCCATCAGAGGACCTGTGTATATGTA 2940
Db 2881 ATACTAAAACTTGCAGAAACAGTGTGAAGTTCTTTCCATCAGAGGACCTGTGTATATGTA 2940
Qy 2941 CTTGGGCTCATAGCTAAACCAAAACAAGGCTGTGATATTTTAAATGTCACAACTGGGAT 3000
Db 2941 CTTGGGCTCATAGCTAAACCAAAACAAGGCTGTGATATTTTAAATGTCACAACTGGGAT 3000
Qy 3001 GCTGTGAGGCATAGTCGCAACATCTGTGGCCAGTGGTTCAGATGATGTGGAAACAATC 3060
Db 3001 GCTGTGAGGCATAGTCGCAACATCTGTGGCCAGTGGTTCAGATGATGTGGAAACAATC 3060
Qy 3061 TGTAAATGAATTTTCTATCTATCCCAAGCACTCTAAAGTTTGAATCTCGAGTCAACCAAGTCT 3120
Db 3061 TGTAAATGAATTTTCTATCTATCCCAAGCACTCTAAAGTTTGAATCTCGAGTCAACCAAGTCT 3120
Qy 3121 AGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCAATTTGGAGGATGACCGG 3180
Db 3121 AGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCAATTTGGAGGATGACCGG 3180
Qy 3181 TTTGGCAGCAGCTCTACTAGTACATTTTCTGATATCAATGAAGTATACAGGCCAACA 3240
Db 3181 TTTGGCAGCAGCTCTACTAGTACATTTTCTGATATCAATGAAGTATACAGGCCAACA 3240
Qy 3241 TTTTATGACCGATCTGGACCCATAAAGATATAAATTCATTTCCCTTTCTTGTCTCTAGT 3300
Db 3241 TTTTATGACCGATCTGGACCCATAAAGATATAAATTCATTTCCCTTTCTTGTCTCTAGT 3300
Qy 3301 AAACTTGTGAAGATCTGATCTTAAATTCGCTTACTTTGCTTAAACCAAAAAACATCGTAGT 3360
Db 3301 AAACTTGTGAAGATCTGATCTTAAATTCGCTTACTTTGCTTAAACCAAAAAACATCGTAGT 3360
Qy 3361 AGCAGTGATCAAAAGGAGGAAATTTATCTGAAAGTAAAGCAAGCAAGCGCAATC 3420
Db 3361 AGCAGTGATCAAAAGGAGGAAATTTATCTCTGAAAGTAAAGCAAGCAAGCGCAATC 3420

Qy 3421 AGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGAATGATTTTACACCATATCC 3480
Db 3421 AGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGAATGATTTTACACCATATCC 3480
Qy 3481 ACTGTACAGAAAACATTTACAAATTAGAGACTTCTTTATGGGAATAGCAATTTGAAGAC 3540
Db 3481 ACTGTACAGAAAACATTTACAAATTAGAGACTTCTTTATGGGAATAGCAATTTGAAGAC 3540
Qy 3541 ACTGCTAGTACACCAAGCATTTGGAGAAAATGACTTAAATTTCAACAAGAAATTTTGGGTACA 3600
Db 3541 ACTGCTAGTACACCAAGCATTTGGAGAAAATGACTTAAATTTCAACAAGAAATTTTGGGTACA 3600
Qy 3601 GAGAAATCACAGAGAAAATAACAAGCCGAGAGAGGTTAGTAGTGAAGTTCAACAGAGTCA 3660
Db 3601 GAGAAATCACAGAGAAAATAACAAGCCGAGAGAGGTTAGTAGTGAAGTTCAACAGAGTCA 3660
Qy 3661 CATATGAAGATACGTAAGCAAGTTTCAATACAGACACTACCAAGTGGCATTAAGTTCA 3720
Db 3661 CATATGAAGATACGTAAGCAAGTTTCAATACAGACACTACCAAGTGGCATTAAGTTCA 3720
Qy 3721 ATGAGCTCAAGTCTCTCACGAGAGACAGTAGGTAGTGTCTACAACTATGGAACACAGAC 3780
Db 3721 ATGAGCTCAAGTCTCTCACGAGAGACAGTAGGTAGTGTCTACAACTATGGAACACAGAC 3780
Qy 3781 TGTGGAAGCATGATGACTGTGTAAGTACTAAAACTATTAAAGCAAGCCACTATTGTAAG 3840
Db 3781 TGTGGAAGCATGATGACTGTGTAAGTACTAAAACTATTAAAGCAAGCCACTATTGTAAG 3840
Qy 3841 CCACAGCTTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCTCCAGGT 3900
Db 3841 CCACAGCTTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCTCCAGGT 3900
Qy 3901 TCTTCTCATACGCTTCTAGAGAGACAGTCCCTTAAAGCACCCCTCTATTGCTTACAAAT 3960
Db 3901 TCTTCTCATACGCTTCTAGAGAGACAGTCCCTTAAAGCACCCCTCTATTGCTTACAAAT 3960
Qy 3961 AAAAGCTACGAGATTTGATTTAGTTTACAAAGTTCTAGAGATGCTTTTGGCTATGCT 4020
Db 3961 AAAAGCTACGAGATTTGATTTAGTTTACAAAGTTCTAGAGATGCTTTTGGCTATGCT 4020
Qy 4021 ACATGAAAAGACTACAGCAACAAAGAAATGTCATCTTATCTCACTCTGAAAGCTTTG 4080
Db 4021 ACATGAAAAGACTACAGCAACAAAGAAATGTCATCTTATCTCACTCTGAAAGCTTTG 4080
Qy 4081 GCATCTCCAGCAAAAGATGCTATTATTTACTGATACCATCACCATGAAAGGCCAACAGTTT 4140
Db 4081 GCATCTCCAGCAAAAGATGCTATTATTTACTGATACCATCACCATGAAAGGCCAACAGTTT 4140
Qy 4141 GAGTCCAGATTTAACCAAGCAGGTTCTAGAAAGCCCTTAAAGTTATGATCATATTAGATAAA 4200
Db 4141 GAGTCCAGATTTAACCAAGCAGGTTCTAGAAAGCCCTTAAAGTTATGATCATATTAGATAAA 4200
Qy 4201 GAAGATTTATGAGTCTTATTAATCAAAATAACCTGCAACGATCTTCTCAGTCCGCTCC 4260
Db 4201 GAAGATTTATGAGTCTTATTAATCAAAATAACCTGCAACGATCTTCTCAGTCCGCTCC 4260
Qy 4261 ATGGTGTCCAGTCCCATATGGGGTTTCAAGATGATTTACATTTGGTCTTCTCTCCCGTG 4320
Db 4261 ATGGTGTCCAGTCCCATATGGGGTTTCAAGATGATTTACATTTGGTCTTCTCTCCCGTG 4320
Qy 4321 GATATAATGATATTTTCCAGGTAAAGGATATTTCCCTTATTTTTCAGACAAAACATACCA 4380
Db 4321 GATATAATGATATTTTCCAGGTAAAGGATATTTCCCTTATTTTCAGACAAAACATACCA 4380
Qy 4381 CCAATGATGATCGAGGTGCAAGAGCAATTTGCCCATGATGACGAGGTCTTCCATCTGGA 4440
Db 4381 CCAATGATGATCGAGGTGCAAGAGCAATTTGCCCATGATGACGAGGTCTTCCATCTGGA 4440
Qy 4441 ACTGGAGGTCTTGTAAAAAATTTCTTTTCACTTGTGTCAGACGAGCATGAGTCTTACGGAA 4500
Db 4441 ACTGGAGGTCTTGTAAAAAATTTCTTTTCACTTGTGTCAGACGAGCATGAGTCTTACGGAA 4500

4501 ATAAATGAATTCATTCATTCAGATGCTCTCTGTTTTTAAAGTACAGAGACACTGGA 4560
Db ATAAATGAATTCATTCATTCAGATGCTCTCTGTTTTTAAAGTACAGAGACACTGGA 4560
4561 CTACAGGAACATACAGATGATTAAGTCTGCTGTTTAAAGTACAGAGACACTGGA 4620
Db CTACAGGAACATACAGATGATTAAGTCTGCTGTTTAAAGTACAGAGACACTGGA 4620
4621 CAGCCAGCAACCACTGAGTGAATATGATGATTCATTCAGATTCCTCAAGATATTCATAT 4680
Db CAGCCAGCAACCACTGAGTGAATATGATGATTCATTCAGATTCCTCAAGATATTCATAT 4680
4681 TCTGATTTGTTGAGCAGATATCCATTAATCTTTAGAAAGTGTTCCTCTAAGTTTTTCG 4740
Db TCTGATTTGTTGAGCAGATATCCATTAATCTTTAGAAAGTGTTCCTCTAAGTTTTTCG 4740
4741 GGGATTTCTGGATGAGTGGGTTGCTCAAGAAAGCTCAGCTAGCAGCACCACAAAGC 4800
Db GGGATTTCTGGATGAGTGGGTTGCTCAAGAAAGCTCAGCTAGCAGCACCACAAAGC 4800
4801 ACAGATTTGTTACTAGTGTAAACCAATTCAGATGATACACCAATGTGCCGTATACCTC 4860
Db ACAGATTTGTTACTAGTGTAAACCAATTCAGATGATACACCAATGTGCCGTATACCTC 4860
4861 CTTTCGCAAGAAAGTCTTAAGATTAGTCAATTAATTTGAGTAGTTCACTCAACTAAATGT 4920
Db CTTTCGCAAGAAAGTCTTAAGATTAGTCAATTAATTTGAGTAGTTCACTCAACTAAATGT 4920
4921 CATGAGACTGGGTTTTTAAACATTAAGGAGAGATATCCTCAACATTTGATGACATATGC 4980
Db CATGAGACTGGGTTTTTAAACATTAAGGAGAGATATCCTCAACATTTGATGACATATGC 4980
4981 CTTTACTCTGAGTTTCCCATTTGCTGTCACACTGCACATTCAGACTTCCTGTCGGAGG 5040
Db CTTTACTCTGAGTTTCCCATTTGCTGTCACACTGCACATTCAGACTTCCTGTCGGAGG 5040
5041 TTTCATCAAGAAATTTTCAAGATGTACAGTTTCTCAAAATGATGAAGAGCAGAGGCT 5100
Db TTTCATCAAGAAATTTTCAAGATGTACAGTTTCTCAAAATGATGAAGAGCAGAGGCT 5100
5101 GTGTTGGGAAACACCAAGCACTATGATGATGATGATGATGATGATGATGATGATGAT 5160
Db GTGTTGGGAAACACCAAGCACTATGATGATGATGATGATGATGATGATGATGATGAT 5160
5161 ATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
Db ATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
5221 TCAGAAATATACGACTGACTAGGAGCAAGAGAGAGATGATGATGATGATGATGATGAT 5280
Db TCAGAAATATACGACTGACTAGGAGCAAGAGAGAGATGATGATGATGATGATGATGAT 5280
5281 GTTACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5340
Db GTTACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5340
5341 TCTCTTTGTTGAGCTTTTGGGGGTTTCATGATTCATTTACCAAGTTTAAAGTTTCACT 5400
Db TCTCTTTGTTGAGCTTTTGGGGGTTTCATGATTCATTTACCAAGTTTAAAGTTTCACT 5400
5401 TACCATTGTATGCAAGAGCAAGCACTGAATATCTACATAGGTTTCTATTTTCTTTTAT 5460
Db TACCATTGTATGCAAGAGCAAGCACTGAATATCTACATAGGTTTCTATTTTCTTTTAT 5460
5461 TTTAAAGCGTAAATGACAGTGAACAAATATGAGGATATGAGAGCACCCTTCACAAGTT 5520
Db TTTAAAGCGTAAATGACAGTGAACAAATATGAGGATATGAGAGCACCCTTCACAAGTT 5520
5521 ATTTCTGTAATGATTTTAGGTTAAATATACAGATGCTGTTGATGTTAACTAACTGTTGA 5580
Db ATTTCTGTAATGATTTTAGGTTAAATATACAGATGCTGTTGATGTTAACTAACTGTTGA 5580
5581 AAGCAGGAATCAGTGTCTCTAAGGCTGCATCCTATTATACCAATGGGTTGTGCTATAAC 5640

Db 5581 AAGCAGGAATCAGTGTCTCTAAGGCTGCATCCTATTACCACAATGGGTTGTGCTATAAC 5640
Qy 5641 TGGCTGGTATTAGAGAGGGAAC 5662
Db 5641 TGGCTGGTATTAGAGAGGGAAC 5662

RESULT 2
US-10-782-244-2
; Sequence 2, Application US/10782244
; Publication NO. US20040253677A1
; GENERAL INFORMATION:
; APPLICANT: Sabatini, David M.
; APPLICANT: Kim, Do-Hyung
; APPLICANT: Sarbasov, Dos D.
; TITLE OF INVENTION: mTOR KINASE-ASSOCIATED PROTEINS
; FILE REFERENCE: WIBL-P01-010
; CURRENT APPLICATION NUMBER: US/10/782,244
; CURRENT FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/448,035
; PRIOR FILING DATE: 2003-02-18
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 5127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-782-244-2

Query Match 90.6%; Score 5127; DB 18; Length 5127;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 ATGGCGCGGATCGCGCGCGCGCGCTCTCTGAAAGAACCTCCGAGTACGAGGCGGAATGAC 87
Db 1 ATGGCGCGGATCGCGCGCGCGCGCTCTCTGAAAGAACCTCCGAGTACGAGGCGGAATGAC 60
Qy 88 AGCGCGAGGAGAAAGTCCCGCTGGATCTGACCCGAGAACCTTCTGATAACTTTAAGAGAG 147
Db 61 AGCGCGAGGAGAAAGTCCCGCTGGATCTGACCCGAGAACCTTCTGATAACTTTAAGAGAG 120
Qy 148 ATTCTCCAAAATGTGGCCAGATTGCGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG 207
Db 121 ATTCTCCAAAATGTGGCCAGATTGCGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG 180
Qy 208 AATAACTTTTACTAAGCTTCTTTGTGATATTGGCCACAGTGAAGAAAACCTGGGCTTTTAC 267
Db 181 AATAACTTTTACTAAGCTTCTTTGTGATATTGGCCACAGTGAAGAAAACCTGGGCTTTTAC 240
Qy 268 TATGAGGATATCATTAATTTGTTGGGTTAGCTTTTAAATGAAAGAAAAGAGTGGCA 327
Db 241 TATGAGGATATCATTAATTTGTTGGGTTAGCTTTTAAATGAAAGAAAAGAGTGGCA 300
Qy 328 GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTCCTCAGAAAGGTG 387
Db 301 GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTCCTCAGAAAGGTG 360
Qy 388 CTAAATAATGAAATGAGATTAATTAATAGTAGTGCATTCATCAACAGAGCAACGAG 447
Db 361 CTAAATAATGAAATGAGACTTAATTAATAGTAGTGCATTCATCAACAGAGCAACGAG 420
Qy 448 GTAGAGAGACACAGCACTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCTTGG 507
Db 421 GTAGAGAGACACAGCACTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCTTGG 480
Qy 508 TTTCTCTAGTTCTGTGACCAACTCATTAAATGTCAGTTGGAAATGATGGACTTCAAGAAAGA 567
Db 481 TTTCTCTAGTTCTGTGACCAACTCATTAAATGTCAGTTGGAAATGATGGACTTCAAGAAAGA 540
Qy 568 GACAGATGGTCCGAGCATGCAATGCCATTATCTGTGAACCTAGCACTTCAGAAATCCAGAG 627
Db 541 GACAGATGGTCCGAGCATGCAATGCCATTATCTGTGAACCTAGCACTTCAGAAATCCAGAG 600

|||||
2761 AAAAACTGAAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAAATTTGGGGTCTC 2820
|||||
2848 AATTTGCTACAGAGAAAACGTGATTCAGATATACATAAACTTGCAAAAACAGTGTGAA 2907
|||||
2821 AATTTGCTACAGAGAAAACGTGATTCAGATATACATAAACTTGCAAAAACAGTGTGAA 2880
|||||
2908 GTTCTTTCCATCAGAGGACCTGTGTATATGTATGTACCTTTGGGCTCATAGCTAAACCAACAA 2967
|||||
2881 GTTCTTTCCATCAGAGGACCTGTGTATATGTATGTACCTTTGGGCTCATAGCTAAACCAACAA 2940
|||||
2968 GGCCTGTGATATTTCTAAATGTCAAACTGGGATGTGTGAGGCATAGTGCAGAAACATCTG 3027
|||||
2941 GGCCTGTGATATTTCTAAATGTCAAACTGGGATGTGTGAGGCATAGTGCAGAAACATCTG 3000
|||||
3028 TGGCCAGTGGTTCAGATGATGTGAAACAACTCTGTATGTAACTTTTCATCTATCCCAAGC 3087
|||||
3001 TGGCCAGTGGTTCAGATGATGTGAAACAACTCTGTATGTAACTTTTCATCTATCCCAAGC 3060
|||||
3088 ACTCTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAAATAGTGAAGTGAATCTGTG 3147
|||||
3061 ACTCTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAAATAGTGAAGTGAATCTGTG 3120
|||||
3148 CCATCGAGTATGTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3207
|||||
3121 CCATCGAGTATGTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
|||||
3208 TTCCTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCATTAAG 3267
|||||
3181 TTCCTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCATTAAG 3240
|||||
3268 GATAAAATTCATTCCTTTCTTTGCTCTAGTAAACTTTGTAAGAACTCGTATCTTAAT 3327
|||||
3241 GATAAAATTCATTCCTTTCTTTGCTCTAGTAAACTTTGTAAGAACTCGTATCTTAAT 3300
|||||
3328 TCGCTTACTTTGCTTAAACAAAAACATCGTAGTACAGTGATCCAAAGAGGAGGAAATTA 3387
|||||
3301 TCGCTTACTTTGCTTAAACAAAAACATCGTAGTACAGTGATCCAAAGAGGAGGAAATTA 3360
|||||
3388 TCATCTGAAGTAAAGACAGACAGGCGAATTCAGACACATTTACGAGCCAGTCTTGAT 3447
|||||
3361 TCATCTGAAGTAAAGACAGACAGGCGAATTCAGACACATTTACGAGCCAGTCTTGAT 3420
|||||
3448 TTTAATCATAGTATGATGATTTTACCCATATCCACTGTACAGAAAACATTAACAATTAGAG 3507
|||||
3421 TTTAATCATAGTATGATGATTTTACCCATATCCACTGTACAGAAAACATTAACAATTAGAG 3480
|||||
3508 ACTTCATTTATGGGGAATGAAGCAATTTGAAGACACTGGTAGTACACCAAGCATTTGAGAA 3567
|||||
3481 ACTTCATTTATGGGGAATGAAGCAATTTGAAGACACTGGTAGTACACCAAGCATTTGAGAA 3540
|||||
3568 AATGACTTAAATTCACCAAGAAATTTTGGTACAGAGAAATTCAGAGAAATACAGCCGA 3627
|||||
3541 AATGACTTAAATTCACCAAGAAATTTTGGTACAGAGAAATTCAGAGAAATACAGCCGA 3600
|||||
3628 GAGAGTTAGTATGATGATGATTTTCAACGAGCTCACATATGAAGATAGTAGCCAAAGTTTC 3687
|||||
3601 GAGAGTTAGTATGATGATGATTTTCAACGAGCTCACATATGAAGATAGTAGCCAAAGTTTC 3660
|||||
3688 AATACAGACATCAACAAAGTGGCAATAGTTTCAATGAGTCAAGTCTTCCAGAGAGACA 3747
|||||
3661 AATACAGACATCAACAAAGTGGCAATAGTTTCAATGAGTCAAGTCTTCCAGAGAGACA 3720
|||||
3748 GTAGGTGATAGTCTACATATGAGACAGACTGTGGAGCATGATGATGTTAGT 3807
|||||
3721 GTAGGTGATAGTCTACATATGAGACAGACTGTGGAGCATGATGATGTTAGT 3780
|||||
3808 ACTTAAATCTATTAAAGCAAGCCACTATTTGAGCCCAACAGTCTAAACATCTGCTCTCTCC 3867
|||||
3781 ACTTAAATCTATTAAAGCAAGCCACTATTTGAGCCCAACAGTCTAAACATCTGCTCTCTCC 3840
|||||
3868 AAATCAAATTCGGTGTCCCTGGTGCCTCCAGGTTCTTCTCATACGCTTCTCTAGAGAGCA 3927
|||||

Db
3841 AAATCAAATTCGGTGTCCCTCGCTCCAGGTTCTTCTCATACGCTTCTCTAGAGAGCA 3900
QY
3928 CAGTCCCTTTAAAGCACCTCTATTGTCTACAATTAATAAGCTTAGCAGATTTGTAACCTTTAGT 3987
Db
3901 CAGTCCCTTTAAAGCACCTCTATTGTCTACAATTAATAAGCTTAGCAGATTTGTAACCTTTAGT 3960
QY
3988 TACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAAGACTACAGCAACAAAGA 4047
Db
3961 TACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAAGACTACAGCAACAAAGA 4020
QY
4048 ATGCATCCATCTCTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAAGATGTGCTATTT 4107
Db
4021 ATGCATCCATCTCTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAAGATGTGCTATTT 4080
QY
4108 ACTGATACCATCACCATGAAGCCCAACAGTTTGTAGTCCAGATTAACCAACCAAGCAGTTTC 4167
Db
4081 ACTGATACCATCACCATGAAGCCCAACAGTTTGTAGTCCAGATTAACCAACCAAGCAGTTTC 4140
QY
4168 ATGAAAGCTTTAAGTTATGTCATCATTTAGATAAAGAAAGATTTATTGAGTCCCTATTAAATCAA 4227
Db
4141 ATGAAAGCTTTAAGTTATGTCATCATTTAGATAAAGAAAGATTTATTGAGTCCCTATTAAATCAA 4200
QY
4228 AATACCTTGCAACGATCTTCTCAGTGGCTCCATGGTTCAGTGCACATATCGGGGT 4287
Db
4201 AATACCTTGCAACGATCTTCTCAGTGGCTCCATGGTTCAGTGCACATATCGGGGT 4260
QY
4288 TCAGATGATTAATCATTTGGTCTTGTCTCTCCGGTGGATATAAATGATATATTTCCAGGTAAAG 4347
Db
4261 TCAGATGATTAATCATTTGGTCTTGTCTCTCCGGTGGATATAAATGATATATTTCCAGGTAAAG 4320
QY
4348 GATATTTCCCTATTTTTCAGACAAAAAATACACCAACATGATGATCGAGGTGCAAGAGCA 4407
Db
4321 GATATTTCCCTATTTTTCAGACAAAAAATACACCAACATGATGATCGAGGTGCAAGAGCA 4380
QY
4408 TTTGCCATGATGACAGAGGCTTCCATCTGGAACCTGGAGGCTTTGTAAAAAATTTCTTTT 4467
Db
4381 TTTGCCATGATGACAGAGGCTTCCATCTGGAACCTGGAGGCTTTGTAAAAAATTTCTTTT 4440
QY
4468 CACTTGTACGACAGCAGATGATGCTTACGGAATAATGAATTTCAATCCATTCAGATGCC 4527
Db
4441 CACTTGTACGACAGCAGATGATGCTTACGGAATAATGAATTTCAATCCATTCAGATGCC 4500
QY
4528 TCTCTGTTTTTGAAGAAGTACAGAGACACTGGACTACAGGAACATACAGATGATTAATCTGC 4587
Db
4501 TCTCTGTTTTTGAAGAAGTACAGAGACACTGGACTACAGGAACATACAGATGATTAATCTGC 4560
QY
4588 CTTTATGCTGTGTGTAATAATTTCTGGGTTTTCCAGCCCGCAACCAACTGAGTGCATTA 4647
Db
4561 CTTTATGCTGTGTGTAATAATTTCTGGGTTTTCCAGCCCGCAACCAACTGAGTGCATTA 4620
QY
4648 TGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4707
Db
4621 TGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4680
QY
4708 AATCCCTTTAGAGTGGTTCCTCTTAAGTTTTCGGGATTTTCTGGATGAGTGAATGGGGTG 4767
Db
4681 AATCCCTTTAGAGTGGTTCCTCTTAAGTTTTCGGGATTTTCTGGATGAGTGAATGGGGTG 4740
QY
4768 TCTCAAGAAGGCTCAGCTAGCAGACCAAAAGCAAGAAATTTGTTACTAGTCTTAAAAACA 4827
Db
4741 TCTCAAGAAGGCTCAGCTAGCAGACCAAAAGCAAGAAATTTGTTACTAGTCTTAAAAACA 4800
QY
4828 ATTTCAGATGATACACCAATGTGCCGTATATCTCTTCCGCAAAAGAGTTCTTAAGATTAGTTC 4887
Db
4801 ATTTCAGATGATACACCAATGTGCCGTATATCTCTTCCGCAAAAGAGTTCTTAAGATTAGTTC 4860
QY
4888 ATTAATTTGAGTGTGCTCAAGTTTCAACTAAATGTCAAGAGACTGGGGCTTTTAAACAATTAAAG 4947
Db
4861 ATTAATTTGAGTGTGCTCAAGTTTCAACTAAATGTCAAGAGACTGGGGCTTTTAAACAATTAAAG 4920
QY
4948 GAGAGATATCTCAAAACATTTGATGACATATGCTCTTACTCTGAGGTTTCCCATTTGCTG 5007
Db
4921 GAGAGATATCTCAAAACATTTGATGACATATGCTCTTACTCTGAGGTTTCCCATTTGCTG 4980

```
QY 5008 TCACACTGCACATTCAGACTTCGGTTCGAGGTTTCATACAGAATTAATTTCAAGATGTA 5067
Db 4981 TCACACTGCACATTCAGACTTCGGTTCGAGGTTTCATACAGAATTAATTTCAAGATGTA 5040
QY 5068 CAGTTTCTACAAATGCATGAGAGCAGAGCTGTGTGGCAACACACCAAGCAACCT 5127
Db 5041 CAGTTTCTACAAATGCATGAGAGCAGAGCTGTGTGGCAACACACCAAGCAACCT 5100
QY 5128 ATAGTTGATACATCTGCTGAATCCTGA 5154
Db 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127

RESULT 3
US-09-814-353-19068/c
; Sequence 19068, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19068
; LENGTH: 4657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 105, 4656, 4657
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19068

Query Match 39.6%; Score 2244.2; DB 10; Length 4657;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2386; Conservative 0; Mismatches 3; Indels 100; Gaps 4;

QY 3175 GACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCTTGATATCAATGAAGATACAGAG 3234
Db 4634 GACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCTTGATATCAATGAAGATACAGAG 4575
QY 3235 CCAACATTTTATGACCGATCTGACCCATTAAGGATAAAATTCATTCCTTTCTTGCT 3294
Db 4574 CCAACATTTTATGACCGATCTGACCCATTAAGGATAAAATTCATTCCTTTCTTGCT 4515
QY 3295 TCTAGTAAATCTTGAAGATCTGATCTTAATTCGTTACTTTGCTTAACCAAAAAACAT 3354
Db 4514 TCTAGTAAATCTTGAAGATCTGATCTTAATTCGTTACTTTGCTTAACCAAAAAACAT 4455
QY 3355 CGTAGTAGCAGTGATCCAAAGGAGGGAATTAATCATCTCAAGTAAGACACCAAGCAGG 3414
Db 4454 CGTAGTAGCAGTGATCCAAAGGAGGGAATTAATCATCTCAAGTAAGACACCAAGCAGG 4395
QY 3415 CGAATCAGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGATGATTTTACACCC 3474
```

```
Db 4394 CGAATCAGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGATGATTTTACACCC 4335
QY 3475 ATATCCACTGTACAGAAAACATTTACAAATTAGAGACTTCATTTATGGGGAATTAAGCAAT 3534
Db 4334 ATATCCACTGTACAGAAAACATTTACAAATTAGAGACTTCATTTATGGGGAATTAAGCAAT 4275
QY 3535 GAAGACACTGTGTAGTACACCAAGCATTTGGAGAAAATGACTTAAATTTCAACAAATTTT 3594
Db 4274 GAAGACACTGTGTAGTACACCAAGCATTTGGAGAAAATGACTTAAATTTCAACAAATTT 4215
QY 3595 GGTACAGAGAAATCAGAGAGAAAATACAGCCGAGAGAGGTTAGTAGTAGAGAAAGTTCAACG 3654
Db 4214 GGTACAGAGAAATCAGAGAGAAAATACAGCCGAGAGAGGTTAGTAGTAGAGAAAGTTCAACG 4155
QY 3655 AGCTCAATATGAAGTACGTACCCAAAGTTTCAATACAGACACTACAAAGTGGGATA 3714
Db 4154 AGCTCAATATGAAGTACGTACCCAAAGTTTCAATACAGACACTACAAAGTGGGATA 4095
QY 3715 AGTTCAATGAGCTCAAGTCTCTTCCAGAGAGACAGTAGGTGTAGATGCTACAACCTATGGAC 3774
Db 4094 AGTTCAATGAGCTCAAGTCTCTTCCAGAGAGACAGTAGGTGTAGATGCTACAACCTATGGAC 4035
QY 3775 ACAGACTGTGGAAGCATGAGTACTGTGTTAAGTACTAAACTATTAAGACAAAGCACTAT 3834
Db 4034 ACAGACTGTGGAAGCATGAGTACTGT----- 4009
QY 3835 TTGACGCCACAGCTCAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCT 3894
Db 4008 ----- 4009
QY 3895 CCAGGTTCTTCTCATACGGTTCCTAGAAAGACACAGTCCCTTAAAGCACCCCTCTATTGCT 3954
Db 4008 ---GGTTCTTCTCATAGGCTTCTAGAGAGACACAGTCCCTTAAAGCACCCCTCTATTGCT 3952
QY 3955 ACAATTAAGTCTAGCAGATTGTAACTTTAGTTACAAAGTCTTAGAGATGCTTTTGGC 4014
Db 3951 ACAATTAAGTCTAGCAGATTGTAACTTTAGTTACAAAGTCTTAGAGATGCTTTTGGC 3892
QY 4015 TATGCTTACACTGAAAGACACTACAGCAACAAAGATCCATCCCTTATCTCACTCTGAA 4074
Db 3891 TATGCTTACACTGAAAGACACTACAGCAACAAAGATCCATCCCTTATCTCACTCTGAA 3832
QY 4075 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGAAGGCCAAC 4134
Db 3831 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGAAGGCCAAC 3772
QY 4135 AGTTTGGAGTCCAGATTAAACACCAAGCAGTTTCATGAAAGCCCTTAAGTTATGATCATTA 4194
Db 3771 AGTTTGGAGTCCAGATTAAACACCAAGCAGTTTCATGAAAGCCCTTAAGTTATGATCATTA 3712
QY 4195 GATTAAGAAAGATTTTATTTAGTCTCTTATTAATCAAAATACCTCGCAACGATCTTCTCAGTG 4254
Db 3711 GATTAAGAAAGATTTTATTTAGTCTCTTATTAATCAAAATACCTCGCAACGATCTTCTCAGTG 3652
QY 4255 CGGTCCATCGTGTCCAGTGCCCATATGCGGGTTTCCAGATGATTTACATTTGCTTTGCTCTC 4314
Db 3651 CGGTCCATCGTGTCCAGTGCCCATATGCGGGTTTCCAGATGATTTACATTTGCTTTGCTCTC 3592
QY 4315 CGGTCCGATATAAATGATATATTTCCAGGTAAAGGATATTCCTCTATTTTACAGAAAAAAC 4374
Db 3591 CGGTCCGATATAAATGATATATTTCCAGGTAAAGGATATTCCTCTATTTTACAGAAAAAAC 3532
QY 4375 ATACCAACCATGATGATCGAGGTCAAGAGCAATTTGCCCATGATCGAGAGGCTCTTCCA 4434
Db 3531 ATACCAACCATGATGATCGAGGTCAAGAGCAATTTGCCCATGATCGAGAGGCTCTTCCA 3472
QY 4435 TCTGGAATCTGGAGGCTCTGTAAATAATTTCTTTTCACTTGTCTACGACAGAGATGAGTCTT 4494
Db 3471 TCTGGAATCTGGAGGCTCTGTAAATAATTTCTTTTCACTTGTCTACGACAGAGATGAGTCTT 3412
QY 4495 ACGGAAATTAATGAATTCATTCATTCAGATGCTCTCTGTTTTTGAAGATGATGAGAGAC 4554
```


Qy 3535 GAAGACCTGCTAGTACACCAAGCATTTGGAGAAATGACTTAAATTTCCCAAGAAATTTT 3594
Db 4220 GAAGACACCTGCTAGTACACCAAGCATTTGGAGAAATGACTTAAATTTCCCAAGAAATTTT 4161
Qy 3595 GGTACAGAGATCACAGAGAAATACAAAGCGAGAGAGTTAGTAGAGAAAGTTCAAG 3654
Db 4160 GGTACAGAGATCACAGAGAAATACAAAGCGAGAGAGTTAGTAGAGAAAGTTCAAG 4101
Qy 3655 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCAT 3714
Db 4100 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCAT 4041
Qy 3715 AGTTCAATAGCTCAAGTCTCTTACAGAGACAGTAGGTGTAGTGTCTACAACCTATGGAC 3774
Db 4040 AGTTCAATAGCTCAAGTCTCTTACAGAGACAGTAGGTGTAGTGTCTACAACCTATGGAC 3981
Qy 3775 ACAGACTGTGGAGCATGAGTACTGTGTGAAGTACTAAACTATTAAAGACAGCCACTAT 3834
Db 3980 ACAGACTGTGGAGCATGAGTACTGTGTGAAGTACTAAACTATTAAAGACAGCCACTAT 3955
Qy 3835 TTGACGCCACAGTCTAACCATCTGTCTCTCTCCAAATCAAAATTCGGGTGTCCTGTGCCT 3894
Db 3954 ----- 3955
Qy 3895 CCAGGTTCTTCTCATACGCTTCTTAGAAGACAGTCCCTTAAAGCACCCCTCTATTGCT 3954
Db 3954 ---GGTTCTTCCGTCAGGCTTCTTAGAAGACAGTCCCTTAAAGCACCCCTCTATTGCT 3998
Qy 3955 ACAATTAAGTCTAGCAGATGTAACTTTAGTTAGTACAAAGTCTTAGAGATGCTTTTGGC 4014
Db 3897 ACAATTAAGTCTAGCAGATGTAACTTTAGTTAGTACAAAGTCTTAGAGATGCTTTTGGC 3838
Qy 4015 TATGCTACACTGAAAGACCTACAGCAACAAAGATGCATCCATCTTATCTCACTCTGAA 4074
Db 3837 TATGCTACACTGAAAGACCTACAGCAACAAAGATGCATCCATCTTATCTCACTCTGAA 3778
Qy 4075 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGAAGGCCAAC 4134
Db 3777 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGAAGGCCAAC 3718
Qy 4135 AGTTTGGAGTCCAGATTAACACAGCAGGTTGATGAAGCCCTTAAGTTATGATCATTTA 4194
Db 3717 AGTTTGGAGTCCAGATTAACACAGCAGGTTGATGAAGCCCTTAAGTTATGATCATTTA 3658
Qy 4195 GATAAAGAGATTTATTTAGTCTCTTATTAATCAAAATACCTGCAACGATCTCTCAGTG 4254
Db 3657 GATAAAGAGATTTATTTAGTCTCTTATTAATCAAAATACCTGCAACGATCTCTCAGTG 3598
Qy 4255 CGGTCCATGCTGTCAGTGCCACATATGGGGTTTCAGATGATTAATGCTTGTCTCTC 4314
Db 3597 CGGTCCATGCTGTCAGTGCCACATATGGGGTTTCAGATGATTAATGCTTGTCTCTC 3538
Qy 4315 CCGGTGATATAAATGATATATTTCCAGGTAAGAGATTTCCCTATTATTTTCAGACAAAAC 4374
Db 3537 CCGGTGATATAAATGATATATTTCCAGGTAAGAGATTTCCCTATTATTTTCAGACAAAAC 3478
Qy 4375 ATACCACCATGATGATCGAGTGCAGAGCATTTGCCCATGATGAGAGGCTCTTCCA 4434
Db 3477 ATACCACCATGATGATGAGTGCAGAGCATTTGCCCATGATGAGAGGCTCTTCCA 3418
Qy 4435 TCTGGAACCTGGAGTCTTGTAAAAAATTTCTTTTTCACCTTGCTACGACAGAGATGAGTCTT 4494
Db 3417 TCTGGAACCTGGAGTCTTGTAAAAAATTTCTTTTTCACCTTGCTACGACAGAGATGAGTCTT 3358
Qy 4495 ACCGAAATATGAAATCAATCCATTCAGATGCTCTCTGTTTGTGAAAGTACAGAGAC 4554
Db 3357 ACCGAAATATGAAATCAATCCATTCAGATGCTCTCTGTTTGTGAAAGTACAGAGAC 3298
Qy 4555 ACTGGACTACAGGAACATACAGATGATACCTGCTTTATGCTCTGTTATTTGAAATCTG 4614
Db 3297 ACTGGACTACAGGAACATACAGATGATACCTGCTTTATGCTCTGTTATTTGAAATCTG 3238

Qy 4615 GGTTCAGCCCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTCAAGATATT 4674
Db 3237 GGTTCAGCCCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTCAAGATATT 3178
Qy 4675 CCATATTTCTGATTTGGTGTGAGCAGACTATCCATAATCTTTAGAGTGGTTCCTCTAAG 4734
Db 3177 CCATATTTCTGATTTGGTGTGAGCAGACTATCCATAATCTTTAGAGTGGTTCCTCTAAG 3118
Qy 4735 TTTTCGGGGATTTCTGGATGAGTGGGTGCTCAAGAGGCTCAGCTAGCAGAC 4794
Db 3117 TTTTCGGGGATTTCTGGATGAGTGGGTGCTCAAGAGGCTCAGCTAGCAGAC 3058
Qy 4795 AAAAGCACAGAAATTTTACTAGGTGTTAAAACAATTCAGATGATACCAAAATGCGGT 4854
Db 3057 AAAAGCACAGAAATTTTACTAGGTGTTAAAACAATTCAGATGATACCAAAATGCGGT 2998
Qy 4855 ATACTCTTCGCAAGAGTTCTTAAGTTAGTCAATTAATTTGAGTAGTTCACT 4914
Db 2997 ATACTCTTCGCAAGAGTTCTTAAGTTAGTCAATTAATTTGAGTAGTTCACT 2938
Qy 4915 AAATGTCATGAGACTGGGCTTTTAAACAATTAAGGAGAGTATCTCAAAATTTGATGAC 4974
Db 2937 AAATGTCATGAGACTGGGCTTTTAAACAATTAAGGAGAGTATCTCAAAATTTGATGAC 2878
Qy 4975 ATATGCTTTTACTCTGAGGTTTCCCATTGCTGTCACTGCACTTCAGACTTCCGCT 5034
Db 2877 ATATGCTTTTACTCTGAGGTTTCCCATTGCTGTCACTGCACTTCAGACTTCCGCT 2818
Qy 5035 CGAGGTTCTATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCAATGAAGAACA 5094
Db 2817 CGAGGTTCTATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCAATGAAGAACA 2758
Qy 5095 GAGGCTGTGTGGCAACACCAACCAAGCACTATAGTTGATACATCTGCTGAATCTGA 5154
Db 2757 GAGGCTGTGTGGCAACACCAACCAAGCACTATAGTTGATACATCTGCTGAATCTGA 2698
Qy 5155 CCTCATATTTATGATGGATATAGATACATATATATATTTTATATTTTGGATTTCCCTA 5214
Db 2697 CCTCATATTTATGATGGATATAGATACATATATATATTTTATATTTTGGATTTCCCTA 2638
Qy 5215 AAAGCCTCAGAAATACGACTGACTAGGCGAAGCAAGCAGGAGTATCTTCTGTACACTGT 5274
Db 2637 AAAGCCTCAGAAATACGACTGACTAGGCGAAGCAAGCAGGAGTATCTTCTGTACACTGT 2578
Qy 5275 TCCGCACTTACTGCTACATGAACAGTTGGAACCTGCTGACTTTTCTTAAACCAAACTTC 5334
Db 2577 TCCGCACTTACTGCTACATGAACAGTTGGAACCTGCTGACTTTTCTTAAACCAAACTTC 2518
Qy 5335 CTTCTCTCTTTGTTGAGGCTTTTGGGGGTTTCATGATTCATTTACCACTTTTAAAGT 5394
Db 2517 CTTCTCTCTTTGTTGAGGCTTTTGGGGGTTTCATGATTCATTTACCACTTTTAAAGT 2458
Qy 5395 TTCACTTACATTTGATGCAAGGCAAGCACTGAATACCTAATAGGTTTCTATTTTC 5454
Db 2457 TTCACTTACATTTGATGCAAGGCAAGCACTGAATACCTAATAGGTTTCTATTTTC 2398
Qy 5455 TTTTCATTTTAAAGCCTTAATGACAGTGGAACTAATGGATGATGAGAGACCCCTTCA 5514
Db 2397 TTTTCATTTTAAAGCCTTAATGACAGTGGAACTAATGGATGATGAGAGACCCCTTCA 2338
Qy 5515 CAAAGTTATTTCTGAATGA-TTTTAGGGTAAATATACAGATGCTTGTATGTTAACTAAC 5573
Db 2337 CAAAGTTATTTCTGAATGAATTTTAGGGTAAATATACAGATGCTTGTATGTTAACTAAC 2278
Qy 5574 TTGTGAAAGCAGGAATCAGTGTCTTAAGGCTGCAATCTTATTTACCAATGCGGTTGTG 5633
Db 2277 TTGTGAAAGCAGGAATCAGTGTCTTAAGGCTGCAATCTTATTTACCAATGCGG-TGTG 2219
Qy 5634 CTATATGCTGCTGATTTAGAGGGAAC 5662
Db 2218 CTATAACT-GCTGGTATTAGAGGGAAC 2191

RESULT 5

US-10-357-930-26008/c
; Sequence 26008, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF HUMAN PROSTATE CANCER
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26008
; LENGTH: 4602
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 105, 4602
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-26008

Query Match 39.6%; Score 2239.4; DB 18; Length 4602;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 2383; Conservative 0; Mismatches 6; Indels 100; Gaps 4;
QY 3175 GACCGGTTGGCAGCAGCTACTAGTACATTTTCCCTTGATATCAATGAAGATACAGAG 3234
DB 4580 GACCGGTTGGCAGCAGCTACTAGTACATTTTCCCTTGATATCAATGAAGATACAGAG 4521
QY 3235 CCAACATTTATGACCGATCTGACCCCATAAAGGATAAAATTCATTCCTTTCTTTGCT 3294
DB 4520 CCAACATTTATGACCGATCTGACCCCATAAAGGATAAAATTCATTCCTTTCTTTGCT 4461
QY 3295 TCTAGTAACTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAAACCAAAACAT 3354
DB 4460 TCTAGTAACTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAAACCAAAACAT 4401
QY 3355 CGTAGTAGCAGTGTGCAAGAGGAGAAATTCATCTGAAAGTAAGACAGCAACAGG 3414
DB 4400 CGTAGTAGCAGTGTGCAAGAGGAGAAATTCATCTGAAAGTAAGACAGCAACAGG 4341
QY 3415 CGAATCAGAACACTTACGAGCCAGTGTGATTTTAAATTCATGATGATGATTTTACACCC 3474
DB 4340 CGAATCAGAACACTTACGAGCCAGTGTGATTTTAAATTCATGATGATGATTTTACACCC 4281
QY 3475 ATATCCACTGTACAGAAACATTAACAATTAGAGACTTCATTTATGCGGAATAAGACATT 3534
DB 4280 ATATCCACTGTACAGAAACATTAACAATTAGAGACTTCATTTATGCGGAATAAGACATT 4221
QY 3535 GAAGACACTGTGTAGTACACCAAGCAATGGAGAAATGACTTAAATTCACCAAGAAATTTT 3594
DB 4220 GAAGACACTGTGTAGTACACCAAGCAATGGAGAAATGACTTAAATTCACCAAGAAATTTT 4161
QY 3595 GGTACAGAGATTCACAGAGAAATACAAAGGAGAGGTTAGTAGTAAAGTTCAACG 3654

DB 4160 GGTACAGAGATTCACAGAGAAATACAGCCGAGAGAGTTAGTAGTAAAGTTCAACG 4101
QY 3655 AGCTCACATATGAAGATAGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 3714
DB 4100 AGCTCACATATGAAGATAGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 4041
QY 3715 AGTTCAATGAGCTCAAGTCTTTCAGAGAGACAGTAGGTGTAGTGTCAATATGGAC 3774
DB 4040 AGTTCAATGAGCTCAAGTCTTTCAGAGAGACAGTAGGTGTAGTGTCAATATGGAC 3981
QY 3775 ACAGACTGTGGAAGCATGAGTACTCTGTGTAAGTACTAATACTATTAAAGACAGCCACTAT 3834
DB 3980 ACAGACTGTGGAAGCATGAGTACTCTGTGTAAGTACTAATACTATTAAAGACAGCCACTAT 3955
QY 3835 TTGACGCCACAGTCTAACCATCTGTCTCTCTCAAAATTCGGTGTCCCTGGTGCCT 3894
DB 3954 ----- 3955
QY 3895 CCAGGTTCTTCTCATACGTTCTTAGAAGAGACAGAGTCCCTTAAAGCAGCCCTCTATTGCT 3954
DB 3954 ---GGTTCTTCGTCAGGCTTCTTAGAAGAGACAGTCCCTTAAAGCAGCCCTCTATTGCT 3898
QY 3955 ACAATTTAAAGTCTAGCAGATTGTAACTTTAGTTACACAAAGTTCTAGAGATGCTTTTGGC 4014
DB 3897 ACAATTTAAAGTCTAGCAGATTGTAACTTTAGTTACACAAAGTTCTAGAGATGCTTTTGGC 3838
QY 4015 TATGCTACACTGAAAGACTACAGCAACAAAGATGTCATCTTATCTCACTCTGAA 4074
DB 3837 TATGCTACACTGAAAGACTACAGCAACAAAGATGTCATCTTATCTCACTCTGAA 3778
QY 4075 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTACTGATACATCACCATGAGGCCAAC 4134
DB 3777 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTACTGATACATCACCATGAGGCCAAC 3718
QY 4135 AGTTTGTAGTCCAGATTAAACACCAAGAGGTTTCATGAAAGCCTTAAAGTTATGTCATCTTA 4194
DB 3717 AGTTTGTAGTCCAGATTAAACACCAAGAGGTTTCATGAAAGCCTTAAAGTTATGTCATCTTA 3658
QY 4195 GATAAAGAAATTTATTGAGTCTTAAATCAAAATACCCCTGCAACAGATCTTCTCTCAGT 4254
DB 3657 GATAAAGAAATTTATTGAGTCTTAAATCAAAATACCCCTGCAACAGATCTTCTCTCAGT 3598
QY 4255 CGGTCCATGTGTCCAGTGCCACATATGGGGGTTTCAGATGATTACATTTGCTCTCTC 4314
DB 3597 CGGTCCATGTGTCCAGTGCCACATATGGGGGTTTCAGATGATTACATTTGCTCTCTC 3538
QY 4315 CCGGTGGATATAAATGATATATTCAGGTAAAGGATATTCCTCTATTTTCAGACAAAAAC 4374
DB 3537 CCGGTGGATATAAATGATATATTCAGGTAAAGGATATTCCTCTATTTTCAGACAAAAAC 3478
QY 4375 ATACCAACCATGATGATCGAGGTGCAAGAGCATTTGCCCATGATGAGGAGGCTTCCCA 4434
DB 3477 ATACCAACCATGATGATCGAGGTGCAAGAGCATTTGCCCATGATGAGGAGGCTTCCCA 3418
QY 4435 TCTGAACTGGAGTCTTTGTAATAAATTTCTTTTCACTTTGCTACGACAGCAGATGAGTCTT 4494
DB 3417 TCTGAACTGGAGTCTTTGTAATAAATTTCTTTTCACTTTGCTACGACAGCAGATGAGTCTT 3358
QY 4495 ACGGAAATTAATGAATCAATCCATTCAGATGCTCTCTCTGTTTTTAAAGATGAGAGAC 4554
DB 3357 ACGGAAATTAATGAATCAATCCATTCAGATGCTCTCTCTGTTTTTAAAGATGAGAGAC 3298
QY 4555 ACTGGAATTCAGAGAACATACAGATGATACTGCTCTTTATTTGCTGTGTTGTAATTTCTG 4614
DB 3297 ACTGGAATTCAGAGAACATACAGATGATACTGCTCTTTATTTGCTGTGTTGTAATTTCTG 3238
QY 4615 GGTTCACAGCCAGCAACCAACTGAGTGCATATGTAGTCAATTCAGACTTTTCAAGATATT 4674
DB 3237 GGTTCACAGCCAGCAACCAACTGAGTGCATATGTAGTCAATTCAGACTTTTCAAGATATT 3178
QY 4675 CCAATTTCTGATTTGGTGTGAGCAGACTATCCATTAATCTTTTAAAGTGGTTCCCTCTTAAG 4734

3177	CCATATCTCTGATTCGGTGTGACGACATATCCATTAATCCTTTAGAGTGGTTCCTCTTAAG	3118
4735	TTTTCGGGGATTTCTGGATGCAGTGAATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCACCC	4794
3117	TTTTCGGGGATTTCTCGATGCAGTGAATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCACCC	3058
4795	AAAAGCACAGAAATTTTACTAGTGTGTTAAACAATTCAGATGATACCAATGTGCGCGT	4854
3057	AAAAGCACAGAAATTTTACTAGTGTGTTAAACAATTCAGATGATACCAATGTGCGCGT	2998
4855	ATACTCTCTTCGCAAGAAGTTCCTAAGATTAAGTCAATTAATTTGAGTAGTTCAGTTCCAACT	4914
2997	ATACTCTCTTCGCAAGAAGTTCCTAAGATTAAGTCAATTAATTTGAGTAGTTCAGTTCCAACT	2938
4915	AAATGTCAATGAGACTCGGCTTTTAAACAATTAAGGAGAAGTATCCTCAAAACATTTGATGAC	4974
2937	AAATGTCAATGAGACTCGGCTTTTAAACAATTAAGGAGAAGTATCCTCAAAACATTTGATGAC	2878
4975	ATATGCCCTTTACTCTCGAGGTTTCCCATTTGCTGTGCACACTGCAANTTCAGACTTCGCGTGT	5034
2877	ATATGCCCTTTACTCTCGAGGTTTCCCATTTGCTGTGCACACTGCAANTTCAGACTTCGCGTGT	2818
5035	CGGAGGTTTCATCAAGAATTAATTTCAAGATGTACAAGTCTTCTACAAATGCAATGAAGAAGCA	5094
2817	CGGAGGTTTCATCAAGAATTAATTTCAAGATGTACAAGTCTTCTACAAATGCAATGAAGAAGCA	2758
5095	GAGGCTGTGTGGCAACACCAACCAAGCAACCTTAAGTTGATACATCTGCTGAATCCTGA	5154
2757	GAGGCTGTGTGGCAACACCAACCAAGCAACCTTAAGTTGATACATCTGCTGAATCCTGA	2698
5155	CCTCATATTTATGATGTGATAGATACATACTATATATATTTCTATATTTGGGATTTTCCTA	5214
2697	CCTCATATTTATGATGTGATAGATACATACTATATATATTTCTATATTTGGGATTTTCCTA	2638
5215	AAAGCCTCAGAAAAATACGACTGACTAGGCAGCAAAAGACAGGAGTATCTTCTGTACACTGT	5274
2637	AAAGCCTCAGAAAAATACGACTGACTAGGCAGCAAAAGACAGGAGTATCTTCTGTACACTGT	2578
5275	TCGCGAGTTACTGTGTACATGAACGTTGGGAACCTGTCTGACTTTCCTAACCAAAACAACTTC	5334
2577	TCGCGAGTTACTGTGTACATGAACGTTGGGAACCTGTCTGACTTTCCTAACCAAAACAACTTC	2518
5335	CTTCTCTCTTTGTTGAGCCTTTTGGGGGGTTCAATGATTCATTAACCAAGTTTTAAAGT	5394
2517	CTTCTCTCTTTGTTGAGCCTTTTGGGGGGTTCAATGATTCATTAACCAAGTTTTAAAGT	2458
5395	TTCAAGTTTACCAATTTGATGCAAGAGCCAGCACTGAATACCTACATAGTGTTCCTATTTTC	5454
2457	TTCAAGTTTACCAATTTGATGCAAGAGCCAGCACTGAATACCTACATAGTGTTCCTATTTTC	2398
5455	TTTTCATTTTTAAAGCGTAATGACAGTGGAAACAATTAATGGGATATGCAAGAGCACCTTCA	5514
2397	TTTTCATTTTTAAAGCGTAATGACAGTGGAAACAATTAATGGGATATGCAAGAGCACCTTCA	2338
5515	CAAGTTATTTCTGAATGA-TTTTAGGGTAAATAATACAGATGCTTGTATGTTAACTAAC	5573
2337	CAAGTTATTTCTGAATGAATTTTTAGGGTAAATAATACAGATGCTTGTATGTTAACTAAC	2278
5574	TTGTGGAAGCAGGAATCAGTGTCTCTAAGGCTGCATCCTTAATACCAATGGGGTTGTG	5633
2277	TTGTGGAAGCAGGAATCAGTGTCTCTAAGGCTGCATCCTTAATACCAATGGGG-TGTG	2219
5634	CTATACTGCTGGTATTAGAGAGGGNAAC	5662
2218	CTATAACT-GCTGGTATTAGAGAGGGNAAC	2191

RESULT 6

US-09-814-353-11

US-09-014-333-11
; Sequence 11, Application US/09814353

; Publication No. US20030165831A1

; GENERAL INFORMATION:

APPLICANT: Lee, John

Db	552	GCTTGTGAAAAAATTTT	569	QY	4330	GATATATCCAGTAAAGGATATTCCTATTTTCAGACAAAAACATACCACCATGAT	4389
				Db	434	GATATATTTCAAGGTAAGGGTATTTCTTATTTTAGACAAAAACAT--CCCCNCAITGAT	491
RESULT 7							
US-09-814-353-6409							
; Sequence 6409, Application US/09814353							
; Publication No. US20030165831A1							
; GENERAL INFORMATION:							
; APPLICANT: Lee, John							
; APPLICANT: Thompson, Pamela							
; APPLICANT: Lillie, James							
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR							
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND							
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER							
; FILE REFERENCE: MRI-006B							
; CURRENT APPLICATION NUMBER: US/09/814,353							
; CURRENT FILING DATE: 2001-03-21							
; PRIOR APPLICATION NUMBER: US 60/191,031							
; PRIOR FILING DATE: 2000-03-21							
; PRIOR APPLICATION NUMBER: US 60/207,124							
; PRIOR FILING DATE: 2000-05-25							
; PRIOR APPLICATION NUMBER: US 60/211,940							
; PRIOR FILING DATE: 2000-06-15							
; PRIOR APPLICATION NUMBER: US 60/216,820							
; PRIOR FILING DATE: 2000-07-07							
; PRIOR APPLICATION NUMBER: US 60/220,661							
; PRIOR FILING DATE: 2000-07-25							
; PRIOR APPLICATION NUMBER: US 60/257,672							
; PRIOR FILING DATE: 2000-12-21							
; NUMBER OF SEQ ID NOS: 22037							
; SOFTWARE: FastSeq for Windows Version 4.0							
; SEQ ID NO 6409							
; LENGTH: 570							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
; FEATURE:							
; NAME/KEY: misc feature							
; LOCATION: 5, 368, 383, 423, 445, 485, 500, 538, 542, 549, 558							
; OTHER INFORMATION: n = A,T,C or G							
US-09-814-353-6409							
Query Match							
Best Local Similarity 8.2%; Score 465.2; DB 10; Length 570;							
Pred. No. 5.1e-114;							
Matches 516; Conservative 0; Mismatches 38; Indels 4; Gaps 3;							
QY	3912	GCTTCTAGAGGACAGTCCCTTAAAGCACCCCTCTATTGCTACAAATTAAGGCTACG	3971	QY	4156	CCAAGCAGGTTTCATGAAAGCCCTTAAGTTATGTCATCATTAGATAAAGAGATTATTGAGT	4215
Db	14	GCTTCTAGAGGACAGTCCCTTAAAGCACCCCTCTATTGCTACAAATTAAGGCTACG	73	Db	400	CCTTTTCAGGTTTCATGAAAGCCCTTAAGTTATGTCATCATTAGATAAAGAGATTATTGAGT	341
QY	3972	AGATTGTAACCTTTAGTTACACAAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAG	4031	QY	4216	CCTATTATCAAAAATACCCCTGCAACGATCTTCCTCAGTCGGGTCCATGGTGTCCAGTGCC	4275
Db	74	AGATTGTAACCTTTAGTTACACAAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAG	133	Db	340	CCTATTATCAAAAATACCCCTGCAACGATCTTCCTCAGTCGGGTCCATGGTGTCCAGTGCC	281
QY	4032	ACTACAGCAACAAGAAATCATCCATCTTATCTCACTCTGAAGCTTTTGGCATCTCCAGC	4091	QY	4276	ACATATGGGGTTTCAGATGATTACATGCTCTCCCGGTGGATATAATGATATA	4335
Db	134	ACTACAGCAACAAGAAATCATCCATCTTATCTCACTCTGAAGCTTTTGGCATCTCCAGC	193	Db	280	ACATATGGGGTTTCAGATGATTACATGCTCTCCCGGTGGATATAATGATATA	221
QY	4092	AAAAGATGCTCTATTACTGATACCATACCATGAAGGCCCAACAGTTTTCAGTCCAGATT	4151	QY	4336	TTCCAGGTAAAGGATATTCCCTATTTCAGACAAAAACATACCACCATGATATCGA	4395
Db	194	AAAAGATGCTCTATTACTGATACCATACCATGAAGGCCCAACAGTTTTCAGTCCAGATT	253	Db	220	TTCCAGGTAAAGGATATTCCCTATTTCAGACAAAAACATACCACCATGATATCGA	161
QY	4152	AACACCAAGCAGGTTTCATGAAGCCCTTAAGTTATGTCATCATATAGATAAAGAGATTATT	4211	QY	4396	GGTGCAAGAGCATTTCGCCCATGATCAGAGGCTCTCCCATCTGGAACTGGAGGCTCTTGA	4455
Db	254	AACACCAAGCAGGTTTCATGAAGCCCTTAAGTTATGTCATCATATAGATAAAGAGATTATT	313	Db	160	GGTGCAAGAGCATTTCGCCCATGATCAGAGGCTCTTCCTCTGGAACTGGAGGCTCTTGA	101
QY	4212	GAGTCTCTATTATCAAAATACCCCTGCAA--CGATTTCTTCAGTGGGGTCCATGGTGCCA	4270				
Db	314	GAGTCTCTATTATCAAAATACCCCTGCAA--CGATTTCTTCAGTGGGGTCCATGGTGCCA	373				
QY	4271	GTGCCACATATGGGGTTCAG--ATGATTACATTTGGTCTTCTCCCGTGGATATAAAT	4329				
Db	374	GTGCCACATATGGGGTTCAG--ATGATTACATTTGGTCTTCTCCCGTGGATATAAAT	433				

RESULT 8
US-10-357-930-9378/c
; Sequence 9378, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endegge, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-0078CN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9378
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-9378

Query Match 6.8%; Score 385.6; DB 18; Length 452;
Best Local Similarity 99.0%; Pred. No. 1.2e-92;
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4156 CCAAGCAGGTTTCATGAAAGCCCTTAAGTTATGTCATCATTAGATAAAGAGATTATTGAGT 4215
Db 400 CCTTTTCAGGTTTCATGAAAGCCCTTAAGTTATGTCATCATTAGATAAAGAGATTATTGAGT 341
QY 4216 CCTATTATCAAAAATACCCCTGCAACGATCTTCCTCAGTCGGGTCCATGGTGTCCAGTGCC 4275
Db 340 CCTATTATCAAAAATACCCCTGCAACGATCTTCCTCAGTCGGGTCCATGGTGTCCAGTGCC 281
QY 4276 ACATATGGGGTTTCAGATGATTACATGCTCTCCCGGTGGATATAATGATATA 4335
Db 280 ACATATGGGGTTTCAGATGATTACATGCTCTCCCGGTGGATATAATGATATA 221
QY 4336 TTCCAGGTAAAGGATATTCCCTATTTCAGACAAAAACATACCACCATGATATCGA 4395
Db 220 TTCCAGGTAAAGGATATTCCCTATTTCAGACAAAAACATACCACCATGATATCGA 161
QY 4396 GGTGCAAGAGCATTTCGCCCATGATCAGAGGCTCTCCCATCTGGAACTGGAGGCTCTTGA 4455
Db 160 GGTGCAAGAGCATTTCGCCCATGATCAGAGGCTCTTCCTCTGGAACTGGAGGCTCTTGA 101

RESULT 11

US-10-357-930-209/c
; Sequence 209, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-209

Query Match 5.9%; Score 336.4; DB 18; Length 416;
Best Local Similarity 98.7%; Pred. No. 2e-79;
Matches 370; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 4176 CTTAAGTTATGATCATATAGATAAAGAGATTATTAGTCTCTATTAATC-AAAATACCC 4234
DB 416 CTTAAGTTATGATCATATAGATAAAGAGATTATTAGTCTCTATTAATC-AAAATACCC 357
QY 4235 TCGAACGATCTT-CCTCAGTGGGTCCTCATGGTGTCCAGTGCCACATATGGGGGTTTCAGAT 4293
DB 356 TCGAACGATCTTCCCTCAGTGGGTCCTCATGGTGTCCAGTGCCACATATGGGGGTTTCAGAT 297
QY 4294 GATTACATGGTCTTCTCCCGTGGGATATAAATGATATATCCAGTAAAGGATATT 4353
DB 296 GATTACATGGTCTTCTCCCGTGGGATATAAATGATATATCCAGTAAAGGATATT 237
QY 4354 CCCTATTTTCAGACAAAAACATACACCATGATGATCGAGGTGCAAGCATTTGCC 4413
DB 236 CCCTATTTTCAGACAAAAACATACACCATGATGATCGAGGTGCAAGCATTTGCC 177
QY 4414 CATGATGAGGAGGTCTTCCATCTGGAAGTGGTCTTGTAAAAAATTTCTTTTTCACCTTG 4473
DB 176 CATGATGAGGAGGTCTTCCATCTGGAAGTGGTCTTGTAAAAAATTTCTTTTTCACCTTG 117
QY 4474 CTACGACGACGATGATGTTTACGAAATAATGAATTCATTCATTCAGATGCCTCTC-T 4532
DB 116 CTACGACGACGATGATGTTTACGAAATAATGAATTCATTCATTCAGATGCCTCTC-T 57
QY 4533 GTTTTATGAAGTAC 4547
DB 56 GTTTTATGAAGTAC 42

RESULT 12

US-09-814-353-12794
; Sequence 12794, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12794
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-12794

Query Match 5.9%; Score 334; DB 10; Length 341;
Best Local Similarity 100.0%; Pred. No. 7.9e-79;
Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3912 GCTTCCTAGAGACACAGTCCCTTAAAGACCCCTCTATTGCTACAAATTAAGTCTAGC 3971
DB 8 GCTTCCTAGAGACACAGTCCCTTAAAGACCCCTCTATTGCTACAAATTAAGTCTAGC 67
QY 3972 AGATTGTAACCTTACTACACAAAGTTCTAGAGATGCTTTTGCTATGCTACACTCAAG 4031
DB 68 AGATTGTAACCTTACTACACAAAGTTCTAGAGATGCTTTTGCTATGCTACACTCAAG 127
QY 4032 ACTACAGCAACAAAGATGTCATCCCTTATCTCAGTCTGAAGCTTTGGCATCTCCAGC 4091
DB 128 ACTACAGCAACAAAGATGTCATCCCTTATCTCAGTCTGAAGCTTTGGCATCTCCAGC 187
QY 4092 AAAAGATGCTATTCTGATACCATCACCATGAGGCCAACAGTTTTAGTCCAGATT 4151
DB 188 AAAAGATGCTATTCTGATACCATCACCATGAGGCCAACAGTTTTAGTCCAGATT 247
QY 4152 AACACCAAGCAGGTTTCATGAAAGCCTTAAGTTATGTCATCATTAAGAGAGATTATT 4211
DB 248 AACACCAAGCAGGTTTCATGAAAGCCTTAAGTTATGTCATCATTAAGAGAGATTATT 307
QY 4212 GAGTCTTATTAATCAAAATACCCCTGCAACGATCT 4245
DB 308 GAGTCTTATTAATCAAAATACCCCTGCAACGATCT 341

RESULT 13

US-09-933-797-314/c
; Sequence 314, Application US/09933797
; Patent No. US20020155119A1
; GENERAL INFORMATION:
; APPLICANT: Robert A. Sikes et al.
; TITLE OF INVENTION: Isolation and Use of Fetal Urogenital
; TITLE OF INVENTION: Sinus Expressed Sequences

FILE REFERENCE: 9901-007-999
CURRENT APPLICATION NUMBER: US/09/933,797
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: US/09/482,933
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: PCT/US99/10746
PRIOR FILING DATE: 1999-05/14
PRIOR APPLICATION NUMBER: 60/085,383
PRIOR FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 811
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 314
LENGTH: 450
TYPE: DNA
ORGANISM: Murine
US-09-933-797-314

Query Match 5.7%; Score 321; DB 9; Length 450;
Best Local Similarity 83.4%; Pred. No. 3e-75;
Matches 373; Conservative 4; Mismatches 67; Indels 3; Gaps 1;

QY 2956 AAAACCAAAAGGCTGTGATATCTAAATGTCAAACTGGGATGCTGTGAGGCATAGT 3015
DB 450 AAACCAAGGTAAAGCTGTGATATCTAAATGTCAAACTGGGATGCTGTGAGGCATAGT 3015
QY 3016 CGCAAAACATCTGGCCAGTGGTTCAGATGATGTGGAACCAACTCTGTAAATGAACCTTCA 3075
DB 393 GTGDADGDTCTATGGCCAGTGGTTCAGATGATGTGGAACCAACTCTGTAAATGAACCTTCA 3075
QY 3076 TCTATCCCAAGCATCTAAAGTTGAAGTTCAGATGATGTGGAACCAACTCTGTAAATGAACCTTCA 3135
DB 333 TCTGTCCCAAGCATCTAAAGTTGAAGTTCAGATGATGTGGAACCAACTCTGTAAATGAACCTTCA 3135
QY 3136 AGTGAATCTGTGCCATCGATGATGTTCAATTTGAGGATGACCGGTTTGGCAGCAGCTCT 3195
DB 273 AGTGAATCTGTGCCATCGATGATGTTCAATTTGAGGATGACCGGTTTGGCAGCAGCTCT 3195
QY 3196 ACTAGTACATTTTCTTGTATCAATGAAGATGATGGAACCAACTCTGTAAATGAACCTTCA 3255
DB 213 ACCAGACATTTTCTTGTATCAATGAAGATGATGGAACCAACTCTGTAAATGAACCTTCA 3255
QY 3256 GGACCCATAAGGATAAAATTCATCCCTTCTTGTATCAATGAAGATGATGGAACCAACTCT 3315
DB 153 GGACCCATAAGGATAAAATTCATCCCTTCTTGTATCAATGAAGATGATGGAACCAACTCT 3315
QY 3316 CGTATCTTAAATTCGCTTACTTTGGCTTAAACCAAAACATCGTAGTAGCAGTGTCCAAA 3375
DB 93 CGTATCTTAAATTCGCTTACTTTGGCTTAAACCAAAACATCGTAGTAGCAGTGTCCAAA 3375
QY 3376 GGAGGAAATTTATCATCTGAAAGTAAG 3402
DB 33 GGAGGAAATTTATCATCTGAAAGTAAG 7

RESULT 14
US-09-910-943-73
Sequence 73, Application US/09910943
Patent No. US20020081610A1
GENERAL INFORMATION:
APPLICANT: Hemmati-Brivanlou, Ali
APPLICANT: Altman, Curtis
TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
FILE REFERENCE: 7529/1G148US1
CURRENT APPLICATION NUMBER: US/09/910,943
CURRENT FILING DATE: 2001-07-23
NUMBER OF SEQ ID NOS: 742
SOFTWARE: PatentIn version 3.1
SEQ ID NO 73
LENGTH: 726
TYPE: DNA
ORGANISM: Xenopus laevis
FEATURE:
NAME/KEY: misc_feature

LOCATION: (1)..(726)
OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-73

Query Match 4.9%; Score 280.2; DB 9; Length 726;
Best Local Similarity 69.5%; Pred. No. 4e-64;
Matches 459; Conservative 0; Mismatches 179; Indels 22; Gaps 5;

QY 4369 AAAACACATACACACATGATGATCGAGGTGCAAGACATTTGGCCATGATGACGAGGAT 4428
DB 75 AAGACGGTGCCTCCAGTGAAGAGGAGCAGCTAAATGTTTGGACCTGAAGTAGGAGGC 134
QY 4429 CTTCCATCTGGAACCTGGAGGTCTTTGAAAAATTTCTTTTCACTTGTGTCACACACAGATG 4488
DB 135 ATCCAA-----GGATCCGGGGTGAAGTCCGTTACAGGCGCTTTCGCGTCAGCAGTTC 185
QY 4489 AGTCTTACGGAATTAATGAATTCATTCATTCAGATGCGCTCTCTGTTTGTAGAAATGACA 4548
DB 186 AGCATTTACGGAATTAATGAATTCATTCAGATGCGCTCTCTGTTTGTAGAAATGACA 245
QY 4549 GAAGACATCTGGAATTCAGAGGAACATACAGATGATAAATGCGCTTTATTTGTCTGTATTGAA 4608
DB 246 GAAGACATCTGGAATTCAGAGGAACATACAGATGATAAATGCGCTTTATTTGTCTGTATTGAA 305
QY 4609 ATTCTGGGTTTCCAGCCAGCAACCACTGAGTCAATATGTAGTCAATTCAGACTTTTCAA 4668
DB 306 TTAATGGGTTTACAGCCAGTCAATTCAGTTCATG---AATATAGCCGTTTAGACTTTTCCC 362
QY 4669 GATATTCATATTTCTGATTTGGTGTGAGCAGACTATCCATTAATCTTTTAGAAGTGGT--- 4724
DB 363 GACATTTCCAAATTTCTGATTTGGTGTGAGCAGACTATCCATTAATCTTTTAGAAGTGGT--- 422
QY 4725 --TCCCTCTAAGTTTTCGGGGATTTCTGGATGCAAGTGGGTGTCTCAAGAGGCTCA 4782
DB 423 CATTTCTTCCAGTGTCTGGATATCAGGTTGTAGTATGTTGTGTC---AGGGCTCG 479
QY 4783 GCTAGCAGCACCAAAAGCAGCAATTTTACTAGGTGTTTAAAACAATTTCCAGATGATACA 4842
DB 480 GCAAGCAGTGACAAAAGTACAGAGTTAGTTCTAGGTGGCAATCAATTTCCGAGATACA 539
QY 4843 CCAATGTCCGTATATCTCTTCCAAAGAGTTCTAAGATTAGTCAATTAATTTAGTAGT 4902
DB 540 CTTGTTTTCAGAAATATTTACTTCCGNAAGAGGTTCTTAAGACTTTGTAATTAACCTGAGTAGC 599
QY 4903 TCAGTTTTCACAACTAAATGTCTATGAGACTGGGCTTTTAAACAAATTAAGGAGAGTATCCCAA 4962
DB 600 TCCGTAGGAAACGAAAGCCCAATGAACTGGGCTCTTAACGATTAAGGAGAGTATTTCTCAA 659
QY 4963 ACATT--TGATGACATATGCTTTTACTCTGAGGTTTCCCAATTTTCTGTGTCACACTGCACATT 5021
DB 660 GCCTTTTGTATGACATTTGCGCTNTATTTCTGAGGTTTCCCACTTTATAGCACATTGCACATT 719

RESULT 15
US-09-783-590-8386
Sequence 8386, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 8386
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (204)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (214)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (224)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (226)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (260)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (262)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (315)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (342)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (359)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (360)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (368)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (372)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (375)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (381)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (401)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (404)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (406)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (410)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (417)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (419)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (423)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (425)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (426)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (444)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (445)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (448)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (452)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (457)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (458)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (460)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (465)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (467)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (469)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (470)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (479)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (482)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (489)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-8386
; Query Match 3.3%; Score 186; DB 9; Length 492;
; Best Local Similarity 89.7%; Pred. No. 7.3e-39;
; Matches 253; Conservative 0; Mismatches 21; Indels 8; Gaps 5;
```

```
Qy 5385 TTTTAAGAGTTTCAGTTACCATTCATTCATGCAAGAGCCCAAGCACTGAATACCTACATAGGTT 5444
Db 15 TTTTAAGAGTTTCAGTTACCATTCATTCATGCAAGAGCCCAAGCACTGAATACCTACATAGGTT 74
Qy 5445 TTCTATTTTCTTTTCATTTTAAAGCGTAATGACACAGTGGACAAATAATGGGATATGCAGAA 5504
Db 75 TTCTATTTTCTTTTCATTTTAAAGCGTAATGACACAGTGGACAAATAATGGGATATGCAGAA 134
Qy 5505 GCACCC-TTCACAAAGTTATTTCGAATGA--TTTTAGGGTAAATAATACAGATGCCCTTGT 5561
Db 135 GCACCCTTTCACAAAGTTATTTCGAATGAATTTTAGGGTAAATAATACAGATGCCCTTGT 194
Qy 5562 ATGTTAACTAA--CTTGTGGAAGCAGAGAAATCAGTGTCTCT--AAGGCTGCATCCTATTA 5617
Db 195 TTGTTAACTNAACTTGTGGAAGCAGAGANTNAGTGTCTCTTAAGGCTGGCATCCGATTA 254
Qy 5618 CCACAAATGGGGTTGTGCTATTAACCTGGCTGCTATTAGAGAGGG 5659
Db 255 CCACANTNGGG-TGTGCTATAACTGGCTGGTATTAGGAGAG 295
```

Search completed: April 9, 2005, 15:10:21
Job time : 2048.44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 17:20:47 ; Search time 607.58 Seconds
(without alignments)

15248.348 Million cell updates/sec

Title: US-10-782-244-1

Perfect score: 5662

Sequence: 1 agcgggtgtgactgaacc.....gctggtattagagagggaac 5662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	392.4	6.9	401	US-09-513-999C-942	Sequence 942, App
2	62.4	1.1	447	US-09-270-767-1698	Sequence 1698, Ap
3	62.4	1.1	447	US-09-270-767-16980	Sequence 16980, A
C 4	60	1.1	1141	US-09-806-708B-22	Sequence 22, Appl
5	58.6	1.0	1141	US-09-806-708B-22	Sequence 22, Appl
C 6	53	0.9	7218	US-08-232-463-14	Sequence 14, Appl
7	44	0.8	399	US-09-621-976-8976	Sequence 8976, Ap
C 8	41	0.7	120727	US-09-949-016-15787	Sequence 15787, Ap
9	41	0.7	120727	US-09-949-016-15788	Sequence 15788, A
C 10	40.8	0.7	832	US-09-621-976-2813	Sequence 2813, Ap
11	39.2	0.7	252	US-09-134-000C-532	Sequence 532, App
12	39.2	0.7	3247	US-09-176-664-19	Sequence 19, Appl
13	39.2	0.7	3319	US-09-176-664-17	Sequence 17, Appl
C 14	39	0.7	3001	US-09-539-333D-199	Sequence 199, App
C 15	39	0.7	83665	US-09-949-016-16995	Sequence 16995, A
16	38.8	0.7	2051	US-09-636-215-699	Sequence 699, App
17	38.8	0.7	2051	US-09-685-166A-699	Sequence 699, App
18	38.8	0.7	2051	US-09-679-426-699	Sequence 699, App
19	38.8	0.7	2051	US-09-759-143-699	Sequence 699, App
20	38.8	0.7	2051	US-09-651-236-699	Sequence 699, App
21	38.8	0.7	10478	US-09-445-774-16	Sequence 16, Appl
C 22	38.8	0.7	256171	US-09-949-016-12822	Sequence 12822, A
23	38.8	0.7	256176	US-09-949-016-15524	Sequence 15524, A
C 24	38	0.7	861	US-09-976-594-289	Sequence 289, App
C 25	38	0.7	1896	US-09-107-532A-248	Sequence 248, App
C 26	38	0.7	1664976	US-08-916-421B-1	Sequence 1, Appli
C 27	38	0.7	1664976	US-09-692-570-1	Sequence 1, Appli

28	37.8	0.7	505	4	US-09-621-976-15639	Sequence 15639, A
29	37.8	0.7	1860	4	US-09-919-497-50	Sequence 50, Appl
C 30	37.8	0.7	52494	4	US-09-949-016-16498	Sequence 16498, A
C 31	37.8	0.7	94095	4	US-09-949-016-14389	Sequence 14389, A
32	37.8	0.7	278866	4	US-09-949-016-13922	Sequence 13922, A
33	37.8	0.7	278866	4	US-09-949-016-13923	Sequence 13923, A
34	37.8	0.7	278866	4	US-09-949-016-13924	Sequence 13924, A
35	37.8	0.7	278866	4	US-09-949-016-13925	Sequence 13925, A
36	37.8	0.7	278866	4	US-09-949-016-13926	Sequence 13926, A
37	37.8	0.7	278866	4	US-09-949-016-14699	Sequence 14699, A
38	37.8	0.7	278866	4	US-09-949-016-14700	Sequence 14700, A
39	37.8	0.7	278866	4	US-09-949-016-14701	Sequence 14701, A
40	37.8	0.7	278866	4	US-09-949-016-14702	Sequence 14702, A
41	37.8	0.7	278866	4	US-09-949-016-14703	Sequence 14703, A
42	37.6	0.7	10274	4	US-09-827-688-5	Sequence 5, Appli
C 43	37.6	0.7	90724	4	US-09-949-016-16601	Sequence 16601, A
C 44	37.4	0.7	601	4	US-09-949-016-179262	Sequence 179262, A
C 45	37.4	0.7	263693	4	US-09-949-016-12386	Sequence 12386, A

ALIGNMENTS

RESULT 1

US-09-513-999C-942
; Sequence 942, Application US/09513999C
; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6783961

; FILE REFERENCE: 59 US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C

; PRIOR APPLICATION NUMBER: 2000-02-24

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ IDS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 942

; LENGTH: 401

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 32..400

; US-09-513-999C-942

Query Match 6.9%; Score 392.4; DB 4; Length 401;
Best Local Similarity 99.7%; Pred. No. 6.2e-107;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	4	GGGTTGCTGACTGAAACCCGCTCAATATGGCGCGCATGGCGCGCGCGCTCTCTGAAGAAC 63
Db	8	GTGTTGCTGACTGAAACCCGCTCAATATGGCGCGCATGGCGCGCGCGCTCTCTGAAGAAC 67
QY	64	CTCCGAGTAGCAGGCGCGGAATGACGCGCGGAGAGAACTCCCGCTGGATCTGACCCGA 123
Db	68	CTCCGAGTAGCAGGCGCGGAATGACGCGCGGAGAGAACTCCCGCTGGATCTGACCCGA 127
QY	124	GAACCTTCTGATAAATTAAGAGAGATTCCTCAAAATGTCAGAGATTCGAGGAGATCA 183
Db	128	GAACCTTCTGATAAATTAAGAGAGATTCCTCAAAATGTCAGAGATTCGAGGAGATCA 187
QY	184	AATATGAGAAAGCTAGGCCATCTGAATAAATTTACTAAGCTTCTTGTGATATGGCCAC 243
Db	188	AATATGAGAAAGCTAGGCCATCTGAATAAATTTACTAAGCTTCTTGTGATATGGCCAC 247
QY	244	AGTGAAGAAAAAATGGGCTTTCATATGAGGATATCATAAATTTGTTGGGTTAGCTTTA 303
Db	248	AGTGAAGAAAAAATGGGCTTTCATATGAGGATATCATAAATTTGTTGGGTTAGCTTTA 307


```
QY 2032 CACCCTCATGAGTTAAATGCTGGAATAATGCAAGTGTATTTCAGTGTCTCTTAATCTT 2091
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 977 NMRMYGAAAGKMGCAAMATGEBWADTAGKMCNNNNNNNTDVRMAMKAKUNNNN 918
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2092 TGCTCTTGAAAAACA-AGATCACTTCTCTAAACTTACTGTTCTAGCTGGACTATAG 2150
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 917 AWTACYNRAATNKKATTHMKWTHGAHSKRRHTRTCRTKYNNNNNNARTVYHH 858
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2151 CAGAGATGGATTGGCTAGAGTCACTCTTTCCAAAATTTTAACTGCGAGCTACTGATGCTG 2210
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 857 AARRMNAWTRTNNNNNNNNNACRNTRTWABKSHSCNNNNNNNNNNNNNTWCHYT 798
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2211 CAGACTCTATGCAACAAACATTTAAGGGTATATTGAGAGCTAATGTGAAATCTTTAA 2270
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 797 TANABBCYRANNNAAARMARCTNNYMAAVTTHTDWCYKTWNTWTDWMTWMTBTTT 738
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2271 TAAITGGGGAATTGAGTTGTTAGTACCCAGCTCATGATAAAAACAAACGATTTCTC 2330
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 737 RMTTSTNNNNNNNNNNWACTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 678
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2331 TGAAGCTCTTGATATCTCGATGAAGCATGTGAAGACAAAGCCAACTCTCATGCTCTCAT 2390
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 677 RMTNTKRWYSTTRHHYTCATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 618
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2391 TCAGATGAACACAGGTTATCCACTTGGAGACAAAGGTTTGTCTCTCTGCTGAGATT 2450
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 617 VRKVKVRDATTCTVDVWADSWWYANWRCRDVTVYTRNNYCKSYAHSYVWNSNNAW 558
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2451 TCTCTCCATTCAAAAGATTTCCTATCTCAATGAAGAGAGTTATGTAGCAAAACAATT 2510
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 557 YRRYSARNSMARWTRNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 498
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2511 GGAAGAGTGGCACAGGGAATCAACTCCAAATATGTTGACTTGTATGAGGAACAACCA 2570
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 497 TTVDSMCNNAKSMRGNRRARWAAANDAGAMDHWTYWGNTNNNNNNNNNNNNNN 438
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2571 TGAAGCACTTACTTACCGGAAGCTGTTGATGGTGTAATAGTGTCTCGAGTAA 2630
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 437 CRRAYCNNNNNNACVWHKHQWRWTKYMKAAACNNNNNNNNNNNNNNNNNNNN 378
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2631 CCAAAGATTACAGGCTCTCAGCTCTACCTCCCTATACACCTTTATGAGCAACTAGTACA 2690
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 377 WMMTSDWBHWTVDYTMRAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 318
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2691 CCATAAACAGGCTGCCATTTGTTGGAAGTACAGAAATATTATACAGAACTCTCTGTA 2750
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 YBMAAMSWAAGASBNVYWCWRTYMGKTWNNNNNNNNNNNNNNNNNNNNNNNN 258
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2751 TGTTCTGACACAGATTGGATAAGTGGAAAGAAATTAAGAACTGAAAGCATCTCTTTG 2810
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 AVMTBKRYKYCYAVBYBYBYBYBYBYBYBYBYBYBYBYBYBYBYBYBYBYBY 198
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2811 GGCCTTGGAAATATCGCTCATCA 2835
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 197 YBKWABAVGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 173
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 5

```
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
```

```
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
```

```
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FABI promoters
US-09-806-708B-22
```

Query Match

Best Local Similarity 1.0%; Score 58.6; DB 4; Length 1141;
Matches 85; Conservative 255; Mismatches 335; Indels 9; Gaps 1;

```
QY 1113 GAGCTTTTCAGATGGCTTTGTGGCAGCTGAGGCAAAAACATTTCTCTCCTCGCCAG 1172
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 GMBWKNWSYDVTYVWDDMKRKYRVRVTRGRMENVYVAMBTARHRRYNGWTEAM 271
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1173 ATCCAGGCCACCTCATGATAATATTGGCACTGATCTCTCTGCAATTTATTCGTA 1232
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 272 AYREWTTNNNNNNNAKCKRAKYGNRABVNSTCTTWKSKTTTKVTSWANNCRAGDAN 331
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1233 TGGACTTTTAGAGGCTCTAGTTGAAGTGATAACAAACAGTCATGATCATATCTCAGTTAG 1292
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 KDHKWKWSAAMGVYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 391
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1293 AGCTACCATCTCTTTTAGGAGAGCTTTTACATATGCGCAACACAATTTCTCTCATTCACA 1352
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 392 YKRKTMYNNNGTTMKRMAWYWKMDMBGTNNNNNGRTYYGWTKNKKWYYKWK 451
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1353 TAGCCATCATTTACATGCTTGCACACCTTAATGAATATGGCTGATCTTTGATATCCC 1412
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 452 ANNCWRADHKTCTHNNNTTWKMTYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 511
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1413 CAAGAAAGAGAGCTGCGAGCCAGTGCAGCTTGAACCTGTTTAAACGCTTCCATGAAT 1472
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 512 HANNNDYWKACTWYKYBVCCKMNNYAAWYTKSSNNYTSRYRKTNNNSWRSDTR 571
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1473 GAAGAAACGAGGCTTAAGCTTATAGTCTTATAGACACCATTTATTCAGAAAGCAAT 1532
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 572 SMGPANNYARABHYGKYKNTRWBSHTWBHBRAGAHHYMBMYBAKCHMKWYKAKK 631
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1533 TGAACACACAGAAACGAGATCAGTATCTCCGAGTTCAGAAAGATATATTTATCTTTAA 1592
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 632 YAGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 691
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1593 GGATACAGAGAGCTCTTTTAACTTAACCTTAGAGATAGCAAGTCTCTTCAACATAAGA 1652
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 692 NNNW-----GCWNNATDTRRTMKNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 742
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1653 GAATCTTTGAATGGAAATCTTATAGGACCATTTCTTAAGTGCGCAAAATGTAAATCT 1712
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 743 KAAKHWKRWKAMRGHADAABTTDKENNGAYTKYTTNNNNNTYGVVNTAARDGW 802
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1713 AAGAACTATAAAGATGAACAGTTACACAGTTTGTACAGAGACTCTTTATTTTAA 1772
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 803 ANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 862
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1773 GCCCAGCAGTAAATATATGCCAACCTGGATCTGGATTTTGGCA 1816
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 863 BAYTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 906
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 6

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

```
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
```


OTHER INFORMATION: 1
US-09-134-000C-532

Query Match 0.7%; Score 39.2; DB 4; Length 252;
Best Local Similarity 51.6%; Pred. No. 0.51; Mismatches 0; Gaps 0;
Matches 83; Conservative 0; Indels 78; Indels 0; Gaps 0;
QY 157 AATGTGCCAGATGTCAGGAGATCAATATAGAAAGCTAGGCCATCTGAATACTTT 216
DB 87 AATGGGCTNANTACCGANNAGCTAAATCTAGATTTTNAAGTTGATTAAGATGATTT 146
QY 217 ACTAAGCTCTTTGTGATATGCGCACAGTGAAGAACTGGGCTTTCACTATGAGGAT 276
DB 147 TCTGAGCCCTTTGAAATCCTTCGAAATTAATGAATTTGGTTCTCACCTTAAGCAA 206
QY 277 ATCATATTTGTTTGGGTTAGCTTTATTAATGAACAAA 317
DB 207 AACATGATGTTTGGCGTTCAATTTATCAACTTACGGAAA 247

RESULT 12

US-09-176-664-19
Sequence 19, Application US/09176664
Patent No. 6815539

GENERAL INFORMATION:
APPLICANT: Salkoff, Lawrence
APPLICANT: Schreiber, Matthew
APPLICANT: Silvia, Chris
APPLICANT: The Washington University
APPLICANT: ICAGEN Inc.

TITLE OF INVENTION: A pH Sensitive Potassium Channel in Spermatoocytes

FILE REFERENCE: 018512-000120US
CURRENT APPLICATION NUMBER: US/09/176,664
CURRENT FILING DATE: 1998-10-21
EARLIER APPLICATION NUMBER: US 60/063,138
EARLIER FILING DATE: 1997-10-22
EARLIER APPLICATION NUMBER: US 60/076,172
EARLIER FILING DATE: 1998-02-27
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 19

LENGTH: 3247
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: human slo3-2 (hslo3-2)
US-09-176-664-19

Query Match 0.7%; Score 39.2; DB 4; Length 3247;
Best Local Similarity 49.5%; Pred. No. 3.3;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 3786 AAGCATGAGTACTGTGTAAGTACTAAACTATTAGACAAAGCCACTATTGACGCCACA 3845
DB 2349 AAGCAACAGACTTTGGTAGACAGAGCCCATCATGGAACCCCTCACATCGGATCCTT 2408
QY 3846 GTCTAACCATCTGTCTCTCCAAATCAAAATCGGTGTCCTCGCTCCAGGTTCTTC 3905
DB 2409 GCAAAATTGACTCCTCTCTGACCCGTCACCTCAGTGTCAGAGGAGACTCCAGGTTAC 2468
QY 3906 TCATACGCTTCTAGAGAGCAGTCCCTTAAAGCACCTCTATTGCTACAAATTAAG 3965
DB 2469 AAATGGACATAATGAGAAATCAAACTGCGGAAAGTCCCTATCTTACTGAAGTAA 2528
QY 3966 TCTAGCAGATTGTAACCTTAGTTA 3989
DB 2529 TCCTTCCACATTCACCTTATTGA 2552

RESULT 13

US-09-176-664-17
Sequence 17, Application US/09176664
Patent No. 6815539

GENERAL INFORMATION:
APPLICANT: Salkoff, Lawrence
APPLICANT: Schreiber, Matthew
APPLICANT: Silvia, Chris
APPLICANT: The Washington University
APPLICANT: ICAGEN Inc.

TITLE OF INVENTION: A pH Sensitive Potassium Channel in Spermatoocytes
FILE REFERENCE: 018512-000120US
CURRENT APPLICATION NUMBER: US/09/176,664
CURRENT FILING DATE: 1998-10-21
EARLIER APPLICATION NUMBER: US 60/063,138
EARLIER FILING DATE: 1997-10-22
EARLIER APPLICATION NUMBER: US 60/076,172
EARLIER FILING DATE: 1998-02-27
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 17

LENGTH: 3319
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: human slo3-1 (hslo3-1)
US-09-176-664-17

Query Match 0.7%; Score 39.2; DB 4; Length 3319;
Best Local Similarity 49.5%; Pred. No. 3.4; Mismatches 0; Gaps 0;
Matches 101; Conservative 0; Indels 103; Indels 0; Gaps 0;

QY 3786 AAGCATGAGTACTGTGTAAGTACTAAACTATTAGACAAAGCCACTATTGACGCCACA 3845
DB 2421 AAGCAACAGACTTTGGTAGACAGAGCCCATCATGGAACCCCTCACATCGGATCCTT 2480
QY 3846 GTCTAACCATCTGTCTCTCCAAATCAAAATCGGTGTCCTCGCTCCAGGTTCTTC 3905
DB 2481 GCAAAATTGACTCCTCTCTGACCCGTCACCTCAGTGTCAGAGGAGACTCCAGGTTAC 2540
QY 3906 TCATACGCTTCTAGAGAGCAGTCCCTTAAAGCACCTCTATTGCTACAAATTAAG 3965
DB 2541 AAATGGACATAATGAGAAATCAAACTGCGGAAAGTCCCTATCTTACTGAAGTAA 2600
QY 3966 TCTAGCAGATTGTAACCTTAGTTA 3989
DB 2601 TCCTTCCACATTCACCTTATTGA 2624

RESULT 14

US-09-539-333D-199/c
Sequence 199, Application US/09539333D
Patent No. 6476208

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Bihain, Bernard
APPLICANT: Essioux, Laurent

TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
FILE REFERENCE: GENSET.047AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
CURRENT FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US 60/126,903
PRIOR FILING DATE: 1999-03-30
PRIOR APPLICATION NUMBER: US 60/131,971
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: US 60/132,065
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: US 60/143,928
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: US 60/145,915
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: US 60/146,453
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: US 60/146,452


```

; SEQ ID NO 16995
; LENGTH: 83665
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16995

Query Match          0.7%; Score 39; DB 4; Length 83665;
Best Local Similarity 46.8%; Pred. No. 42;
Matches 123; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 4882 TTAGTCATTAAATTGAGTAGTTTCAGTTCCTCAACTAAATGTCATGAGACTGGGCTTTTAAACA 4941
      |||||
Db 19461 TTTGATTTTATTGTGATAATGATCATTTCTTCAATATTGGATGTGATGCTGCTTGCTGTTT 19402

QY 4942 ATTAAGGAGAAGTATCCTCAACATTTGATGACATATGCGCTTTACTCTGAGGTTTCCCAT 5001
      |||||
Db 19401 ATTTTAGATCAATTTTCTTCAATAATTGAGTCCATATCACATTTCCCTAAAGAGTTTTTCATT 19342

QY 5002 TTGCTGTACACGTGCACATTCAGACTTCCGTCGTCGGAGGTTTCATACAGAAATTATTTCAA 5061
      |||||
Db 19341 ATTTAAACAAGCTCTCCATTTAAATTTCTATTAGATTTTTTAAGCCTCTAAAAAGTAACCCCTA 19282

QY 5062 GATGTACAGTTTTCTCAAAATGCATGAAGAAGCAGAGGCTGTGTTGGCCAAACACCACCAAG 5121
      |||||
Db 19281 GAATTTAAATTAGATGTACTTCTGATATACCATCTAATATAAATTCAAACATTTTACATG 19222

QY 5122 CAACCTATAGTTGATACATCTGC 5144

Db 19221 CTACAAATATTTGCTAGAGGTAC 19199

```

Search completed: April 9, 2005, 13:48:34
Job time : 613.58 secs

```

Query Match          0.7%; Score 39; DB 4; Length 3001;
Best Local Similarity 51.4%; Pred. No. 3.6;
Matches 90; Conservative 85; Indels 0; Gaps 0;

QY      5127 TATAGTTGATACATCTCGTGAATCCTGCACCTCATATTATTTATGGATATAGATACATACT 5186
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1094 TATATATAATATATATGCTATATCTATATACTATATAATATGTTTTCTCATATATATAAAT 1035

QY      5187 ATATPATATTCATATTTGTGGATTTCTCTAAAGCCCTCAGAAAATACGACTGCTAGCACG 5246
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1034 ATATATATATTTGTATTTTAAAAATTAAGTGCTCTTACTAAATATCTGCTGGGCCTT 975

QY      5247 AAAGACAGGAGTATCTTCTGTACACTGTTCCGCAGTTACTGGTACATGACAGTT 5301
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       974 ACAAGCAAAATTTATAGTAAATCTAGGTCCTGAATAGTCTCGTTTTCTGATGACTT 920


RESULT 15
US-09-949-016-16995/c
; Sequence 16995, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; CURRENT APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; CURRENT APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 11:22:35 ; Search time 1884 Seconds
(without alignments)
17790.649 Million cell updates/sec

Title: US-10-782-244-1
Perfect score: 5662
Sequence: 1 agcgggttgactgaacc.....gctggtattagagagggaac 5662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Desc04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002as:*

7: Geneseq2002bs:*

8: Geneseq2003as:*

9: Geneseq2003bs:*

10: Geneseq2003cs:*

11: Geneseq2003ds:*

12: Geneseq2004as:*

13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5662	100.0	5662	13	ADR44434
2	5127	90.6	5127	13	ADR44435
3	4271.4	75.4	8232	10	ADR44435 Human p20
4	3829	67.6	3829	13	ADR44435 Human p20
5	2276.4	40.2	2659	10	ADR44435 Human p20
6	2244.2	39.6	4657	5	ADR44435 Human p20
7	2239.4	39.6	4601	5	ADR44435 Human p20
8	2239.4	39.6	4602	5	ADR44435 Human p20
9	976.8	17.3	1334	10	ADR44435 Human p20
10	874.8	15.5	3129	13	ADR44435 Human p20
11	809.6	14.3	4726	13	ADR44435 Human p20
12	744	13.1	1292	5	ADR44435 Human p20
13	612.8	10.8	765	5	ADR44435 Human p20
14	537.8	9.5	1185	5	ADR44435 Human p20
15	537.8	9.5	1185	5	ADR44435 Human p20
16	537.8	9.5	1185	5	ADR44435 Human p20
17	455.2	8.2	570	5	ADR44435 Human p20
18	455.2	8.2	570	5	ADR44435 Human p20
19	459	8.1	473	3	ADR44435 Human p20
20	392.4	6.9	401	3	ADR44435 Human p20

C	21	385.6	6.8	452	5	ABV09387	Abv09387 Human pro
C	22	383.4	6.8	427	5	ABV39532	Abv39532 Human pro
C	23	383.4	6.8	427	5	ABV30563	Abv30563 Human pro
C	24	357	6.3	770	4	AAI96374	Aai96374 Human neu
C	25	340.8	6.0	1347	5	AAI96374	Aai96374 Human neu
C	26	336.4	5.9	416	5	ABV00218	Abv00218 Human pro
C	27	334	5.9	341	5	ADL38904	Adl38904 Human ova
C	28	321	5.7	450	3	ADP56995	Adp56995 Urogenita
C	29	280.2	4.9	726	6	ABV76822	Abv76822 Frog embr
C	30	166.8	2.9	567	12	ACH75616	Ach75616 Human gen
C	31	163.6	2.9	170	12	ACH89342	Ach89342 Human gen
C	32	133.8	2.4	5769	4	ABL13629	Abli13629 Drosophil
C	33	133.8	2.4	5769	13	ADQ89657	Adq89657 Antagonis
C	34	103.8	1.8	3982	12	ADQ35582	Ado35582 Novel mou
C	35	81	1.4	9139	4	ABL13628	Abli13628 Drosophil
C	36	77.4	1.4	289	5	ADL44180	Adl44180 Human ova
C	37	73	1.3	185	5	ADL63388	Adl63388 Human ova
C	38	57	1.0	185	5	ADL37796	Adl37796 Human ova
C	39	57	1.0	185	5	ADL72657	Adl72657 Human ova
C	40	53.2	0.9	551	6	ABQ20803	Abq20803 Oligonucl
C	41	53.2	0.9	551	6	ABQ20802	Abq20802 Oligonucl
C	42	51.4	0.9	551	6	ABQ20804	Abq20804 Oligonucl
C	43	51.4	0.9	551	6	ABQ20805	Abq20805 Oligonucl
C	44	49	0.9	2000	8	ADA71938	Ada71938 Rice gene
C	45	48.2	0.9	4590	5	AAH24065	Aah24065 Yeast AOD

ALIGNMENTS

RESULT 1

ADR44434

ID ADR44434 standard; cdna; 5662 BP.

XX

AC ADR44434;

XX

DT 18-NOV-2004 (first entry)

XX

DE Human p200 encoding cdna SEQ ID NO:1.

XX

KW mTOR-associated protein; mTOR-AP; cytotostatic; antidiabetic; gene therapy;

KW mTOR-AP modulation; cancer; tumour; diabetes; human; p200; gene; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 28..5154

FT /*tag= a

FT /product= "p200"

XX

PN WO2004074448-A2.

XX

PD 02-SEP-2004.

XX

PF 18-FEB-2004; 2004WO-US004821.

XX

PR 18-FEB-2003; 2003US-0448035P.

XX

PA (WHEH) WHITEHEAD INST BIOMEDICAL RES.

XX

PI Sabatini DM, Kim D, Sarbasov DD.

XX

DR WPI; 2004-635565/61.

DR P-PSDB; ADR44436.

XX

PT New mTOR-associated protein (mTOR-AP) polypeptide, useful for treating or preventing a disorder that is responsive to mTOR-AP modulation, e.g. cancer or diabetes.

XX

PS Claim 14; SEQ ID NO 1; 91pp; English.

XX

CC The present invention describes an isolated mTOR-associated protein (mTOR-AP) polypeptide comprising a sequence that is 70% identical to the 1708

amino acid sequence of SEQ ID NO:3 (ADR44436, P1) or 70% identical to the 327 amino acid sequence of SEQ ID NO:6 (ADR44439, P2), or that is encoded by a nucleic acid that hybridizes under high stringency conditions to a complement sequence of the 351271 nucleic acid sequence of SEQ ID NO:2 (ADR44435, S1) or the 981 nucleic acid sequence of SEQ ID NO:5 (ADR44438, S2). Also described: (1) an isolated antibody, or its fragment, which is specifically immunoreactive with an epitope of an amino acid sequence of P1 or P2; (2) the isolated nucleic acid encoding the polypeptide, or its complement; (3) a vector comprising a nucleic acid sequence encoding an mTOR-AP polypeptide; (4) an isolated host cell comprising the nucleic acid; (5) a method of producing an mTOR-AP polypeptide; (6) a method for detecting the presence of an mTOR-AP polypeptide; (7) a kit for detecting an mTOR-AP polypeptide comprising the antibody and a detectable label for detecting the antibody; (8) a method for detecting the presence of the nucleic acid in a sample; (9) a kit comprising the probe or primer nucleic acid and instructions for use; (10) an isolated, purified or recombinant complex comprising an mTOR polypeptide and an mTOR-AP; (11) a method for identifying a compound which modulates activity of an mTOR-AP polypeptide; (12) a method of inhibiting aberrant activity of an mTOR-AP-expressing cell; (13) a method of treating or preventing a disorder that is responsive to mTOR-AP modulation, in a subject; and (14) a transgenic mouse having germline and somatic cells comprising a chromosomally incorporated transgene that disrupts the genomic mTOR-AP gene and inhibits expression of the gene, where the disruption comprises insertion of a selectable marker sequence resulting in the transgenic mouse exhibiting increased susceptibility to the formation of tumours as compared to the wild type mouse. mTOR-AP sequences cytostatic and antidiabetic activities, and can be used in gene therapy. The polypeptides, polynucleotides, compounds and methods are useful for treating or preventing a disorder that is responsive to mTOR-AP modulation, e.g. cancer or diabetes. The compound, which modulates an mTOR-AP activity or expression is useful for the manufacture of a medicament for treating a disease affected by aberrant mTOR-AP activity or expression. The present sequence encodes human p200, which is an mTOR-AP used in the exemplification of the present invention.

XX SQ Sequence 5662 BP; 1786 A; 1128 C; 1156 G; 1592 T; 0 U; 0 Other;

Query Match 100.0%; Score 5662; DB 13; Length 5662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACGCGGTTGTGACTGAAACCCGTCATATGCGCGGATCGCGCGCGCGCTCTCTGAAG	60
DB	1	ACGCGGTTGTGACTGAAACCCGTCATATGCGCGGATCGCGCGCGCGCTCTCTGAAG	60
QY	61	ACCTCCGAGTACGAGGGGGAATGACACGCGGAGAGAACTCCCGTGTGATCTGACC	120
DB	61	ACCTCCGAGTACGAGGGGGAATGACACGCGGAGAGAACTCCCGTGTGATCTGACC	120
QY	121	CGAGAACCTTCTGTAACTTAAGAGAGATTCTCCAAATGTGGCCAGATTGCGAGGAGTA	180
DB	121	CGAGAACCTTCTGTAACTTAAGAGAGATTCTCCAAATGTGGCCAGATTGCGAGGAGTA	180
QY	181	TCAATATGAGAAAGCTAGGCCATCTGAATACTTTACTAAGCTTCTTTGTGATTTGGC	240
DB	181	TCAATATGAGAAAGCTAGGCCATCTGAATACTTTACTAAGCTTCTTTGTGATTTGGC	240
QY	241	CACAGTGAAGAAACTCGGCTTTCATATGAGGATATCAATTTGTTGGGTTAGCT	300
DB	241	CACAGTGAAGAAACTCGGCTTTCATATGAGGATATCAATTTGTTGGGTTAGCT	300
QY	301	TTATTAAATGAAGCAAAAGAGTGGCAGCAGGCTACGAGCGCTTCGATATCTCATC	360
DB	301	TTATTAAATGAAGCAAAAGAGTGGCAGCAGGCTACGAGCGCTTCGATATCTCATC	360
QY	361	CAAGACTCCAGTATTTCTCCAGAGGTGCTTAAATTTGAAAGTGAATTTAATAGCTAGG	420
DB	361	CAAGACTCCAGTATTTCTCCAGAGGTGCTTAAATTTGAAAGTGAATTTAATAGCTAGG	420
QY	421	TGCATTGACATACACAGAGCAACGAGGTAGAGAGACACAGACTTCGATTAGTACAG	480
DB	421	TGCATTGACATACACAGAGCAACGAGGTAGAGAGACACAGACTTCGATTAGTACAG	480

QY	481	AGATGATTACTGTGAATGCTTCTCTGTTCTCTAGTCTCTGTGACCAACTCATTAATGCA	540
DB	481	AGATGATTACTGTGAATGCTTCTCTGTTCTCTAGTCTCTGTGACCAACTCATTAATGCA	540
QY	541	GTTCGAATATGATGGACTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATGCCATTATC	600
DB	541	GTTCGAATATGATGGACTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATGCCATTATC	600
QY	601	TGTGAATATGATGGACTTCAAGAAATGGTCCGAGCATGCAATGCCATTATC	660
DB	601	TGTGAATATGATGGACTTCAAGAAATGGTCCGAGCATGCAATGCCATTATC	660
QY	661	TTGAAATATGATGGACTTCAAGAAATGGTCCGAGCATGCAATGCCATTATC	720
DB	661	TTGAAATATGATGGACTTCAAGAAATGGTCCGAGCATGCAATGCCATTATC	720
QY	721	TTGCACTTCTTAAATCATCCAAAGACTCGGAGTATGTGCGAGCTGATGAGAAATAGAG	780
DB	721	TTGCACTTCTTAAATCATCCAAAGACTCGGAGTATGTGCGAGCTGATGAGAAATAGAG	780
QY	781	AGAAATTTAGCACTTACTGATTTTCACTACAGACATAGTCCAGATACAGCTGAAGA	840
DB	781	AGAAATTTAGCACTTACTGATTTTCACTACAGACATAGTCCAGATACAGCTGAAGA	840
QY	841	CAGCTCAAGAGACAGAGAGACAGATTTCTAGCCAGTAAATGGGAATCATAGCAACA	900
DB	841	CAGCTCAAGAGAGACAGAGAGACAGATTTCTAGCCAGTAAATGGGAATCATAGCAACA	900
QY	901	TTCCGATCATGGGAGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	960
DB	901	TTCCGATCATGGGAGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	960
QY	961	CTAATAGAGTACTTTGTCATACCAAAATATGGAATTAAGGCGAGGTCTACTTGAAGTCTT	1020
DB	961	CTAATAGAGTACTTTGTCATACCAAAATATGGAATTAAGGCGAGGTCTACTTGAAGTCTT	1020
QY	1021	TATGATATATTTCTGCTTCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1080
DB	1021	TATGATATATTTCTGCTTCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1080
QY	1081	AGTGTAGATCCAGGAGGTTCCTCAAGACAGATTTGAGGCTTTTCAGATGGCTTTTGGCAGCT	1140
DB	1081	AGTGTAGATCCAGGAGGTTCCTCAAGACAGATTTGAGGCTTTTCAGATGGCTTTTGGCAGCT	1140
QY	1141	GAGGCAAAACCTATTTCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1200
DB	1141	GAGGCAAAACCTATTTCTTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1200
QY	1201	TTGGCAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
DB	1201	TTGGCAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
QY	1261	ATAACAAACAGTGTGATCATCTCAGTTAGAGCTACCATCTTTTATAGGAGAGCTTTTA	1320
DB	1261	ATAACAAACAGTGTGATCATCTCAGTTAGAGCTACCATCTTTTATAGGAGAGCTTTTA	1320
QY	1321	CATATGGCAACAACTTCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
DB	1321	CATATGGCAACAACTTCTTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
QY	1381	CTAATGAATATGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
DB	1381	CTAATGAATATGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
QY	1441	GCCTTGAACTGTTTAAACCGCTTCCATGAAATGAAGAAACGAGGACCTTAAGCCTTTATAGT	1500
DB	1441	GCCTTGAACTGTTTAAACCGCTTCCATGAAATGAAGAAACGAGGACCTTAAGCCTTTATAGT	1500
QY	1501	CTTCAATTTAGACCACTTATTTTCAAGAGCAATTTGCAACACACACACACACACACACACACAC	1560
DB	1501	CTTCAATTTAGACCACTTATTTTCAAGAGCAATTTGCAACACACACACACACACACACACACAC	1560

Qy	1561	CTCCGAGTTTCAGAAAGATATATTTATTCCTTTAAGGATACAGAGGAAGCTCTTTTAAATTAAC	1620
Db	1561	CTCCGAGTTTCAGAAAGATATATTTATTCCTTTAAGGATACAGAGGAAGCTCTTTTAAATTAAC	1620
Qy	1621	CTTAGAGATAGCCAAAGTCCTTCAACATAAAGAGAACTTCTGAATGGAATTCGGAATCTTATA	1680
Db	1621	CTTAGAGATAGCCAAAGTCCTTCAACATAAAGAGAACTTCTGAATGGAATTCGGAATCTTATA	1680
Qy	1681	GGGACCAATCTTTAAAGTGGCCAAATGTAAATCTTAAGAAACTATTAAGAGATGAACAGTTACAC	1740
Db	1681	GGGACCAATCTTTAAAGTGGCCAAATGTAAATCTTAAGAAACTATTAAGAGATGAACAGTTACAC	1740
Qy	1741	AGGTTTGTGACGAGCTACTTTATTTTTTCAAGCCAGCAGTAAATATATATGCGCAACCTG	1800
Db	1741	AGGTTTGTGACGAGCTACTTTATTTTTTCAAGCCAGCAGTAAATATATATGCGCAACCTG	1800
Qy	1801	GATCTGGATTTTGCACAGGCCAAGCAGCTCAGCGTGTAGGTTGCGCAGTTTACAGAAATTT	1860
Db	1801	GATCTGGATTTTGCACAGGCCAAGCAGCTCAGCGTGTAGGTTGCGCAGTTTACAGAAATTT	1860
Qy	1861	CTTCTTGAATCTGAAGAGGATGGCCAAAGGCTACTTTAGAAGATCTAGTAAAGGATATGTT	1920
Db	1861	CTTCTTGAATCTGAAGAGGATGGCCAAAGGCTACTTTAGAAGATCTAGTAAAGGATATGTT	1920
Qy	1921	CAGTGGCTCAATGCTTCATCTGGAATGAACCCGAAAGAAAGTCTTCAAAATAATGTTT	1980
Db	1921	CAGTGGCTCAATGCTTCATCTGGAATGAACCCGAAAGAAAGTCTTCAAAATAATGTTT	1980
Qy	1981	TTGACCAACCTTAGTCAACACTACTTTTTTATTTATTTGGAACACTTTCTTGCCACCTCAT	2040
Db	1981	TTGACCAACCTTAGTCAACACTACTTTTTTATTTATTTGGAACACTTTCTTGCCACCTCAT	2040
Qy	2041	GGAGTAAATGCTGGAAAAATGAGTGTATTTTCAGTGTCTCCTTAATCTTTGCTCCTTG	2100
Db	2041	GGAGTAAATGCTGGAAAAATGAGTGTATTTTCAGTGTCTCCTTAATCTTTGCTCCTTG	2100
Qy	2101	AAAAACCAAGATCACTTGCTAAAACTTACTGTTTCTAGCTTGGACTATAGCAGAGATGGA	2160
Db	2101	AAAAACCAAGATCACTTGCTAAAACTTACTGTTTCTAGCTTGGACTATAGCAGAGATGGA	2160
Qy	2161	TTGCTAGAGTCACTCTTCCAAAATTTTAACATGACGCTAGTGCCTGCAGACTCTAT	2220
Db	2161	TTGCTAGAGTCACTCTTCCAAAATTTTAACATGACGCTAGTGCCTGCAGACTCTAT	2220
Qy	2221	GCAACAAAAATTTAAGGATTTATTCGAGAGCTAATGTTGAATCTCTTAATAATTTGGGGA	2280
Db	2221	GCAACAAAAATTTAAGGATTTATTCGAGAGCTAATGTTGAATCTCTTAATAATTTGGGGA	2280
Qy	2281	ATTGAGTTGTTAGTGACCCAGCTACATGATATAAAAAACGAATTTCTCTGAAGCTCTT	2340
Db	2281	ATTGAGTTGTTAGTGACCCAGCTACATGATATAAAAAACGAATTTCTCTGAAGCTCTT	2340
Qy	2341	GATATCTCGATGAAGCATGTGAAGCAAGGCCAATCTTCATGCTCTCATTCAGATGAAA	2400
Db	2341	GATATCTCGATGAAGCATGTGAAGCAAGGCCAATCTTCATGCTCTCATTCAGATGAAA	2400
Qy	2401	CCAGCTTATCCCACTTTGGAGACAAGGTTTGCTTCTCCTGCTGAGATTTCTCTCCATT	2460
Db	2401	CCAGCTTATCCCACTTTGGAGACAAGGTTTGCTTCTCCTGCTGAGATTTCTCTCCATT	2460
Qy	2461	CCAAAAGGATTTTCCCTATCTGAATGAAGAGGTTTATGTAGCAAAACAAATTTGGAAGTGG	2520
Db	2461	CCAAAAGGATTTTCCCTATCTGAATGAAGAGGTTTATGTAGCAAAACAAATTTGGAAGTGG	2520
Qy	2521	CACAGGGAATACAACTCCAAATATGTTGACTTGAATTGAGGAAACAACTCAATGAAGCACTT	2580
Db	2521	CACAGGGAATACAACTCCAAATATGTTGACTTGAATTGAGGAAACAACTCAATGAAGCACTT	2580
Qy	2581	ACTACTTACCGGAAGCCTGTTGATGGTGATAAATCTATGTTGCTGGAGTAAACCAAGATTA	2640
Db	2581	ACTACTTACCGGAAGCCTGTTGATGGTGATAAATCTATGTTGCTGGAGTAAACCAAGATTA	2640
Qy	2641	CAGCGTCTCAGCGTCTACCTGCCTATACACTTTTATGGACAACCTAGTATACCAATAAACCA	2700

Db	2641		CAGCGTCTCACGCTACTACCTGCGCTATACACCTTTATGGACAACCTAGTATACCAATAAACA	2700
Qy	2701	GGCTGCCATTGTTGGTGAAGTACAGAAATATTTATTACAGAACTCTGTGCTAATGTTCTGTACA	2760	
Db	2701	GGCTGCCATTGTTGGTGAAGTACAGAAATATTTATTACAGAACTCTGTGCTAATGTTCTGTACA	2760	
Qy	2761	CCAGATTTGGATAAGTGGGAAGAAATTTAAAAAATCTGAAAGCATCTCTTTTGGGCTTTGGGA	2820	
Db	2761	CCAGATTTGGATAAGTGGGAAGAAATTTAAAAAATCTGAAAGCATCTCTTTTGGGCTTTGGGA	2820	
Qy	2821	AATATCGGCTCATCAAAATTCGGGTCTCAATTTCTGCTACAGGAAGAAACGTTGATTCAGAT	2880	
Db	2821	AATATCGGCTCATCAAAATTCGGGTCTCAATTTCTGCTACAGGAAGAAACGTTGATTCAGAT	2880	
Qy	2881	ATACTAAAACTTCAAAACAGTGTGAAGTCTTTTCCATCAGAGGGACCTGTGTATATGTA	2940	
Db	2881	ATACTAAAACTTCAAAACAGTGTGAAGTCTTTTCCATCAGAGGGACCTGTGTATATGTA	2940	
Qy	2941	CTTGGGCTCATAGCTAAAAACAAACAAAGCTGTGATATCTTAAAAATGTCAAACTGGGAT	3000	
Db	2941	CTTGGGCTCATAGCTAAAAACAAACAAAGCTGTGATATCTTAAAAATGTCAAACTGGGAT	3000	
Qy	3001	GCTGTGAGGCATAGTTCGCAAAACATCTGTGGCCAGTGGTTCAGATGATGTGGAAACAATC	3060	
Db	3001	GCTGTGAGGCATAGTTCGCAAAACATCTGTGGCCAGTGGTTCAGATGATGTGGAAACAATC	3060	
Qy	3061	TGTAATGAACTTTTCATCTATCCCAAGCACCTCAAGTTTGAACCTCGAGTCAACACGCTCT	3120	
Db	3061	TGTAATGAACTTTTCATCTATCCCAAGCACCTCAAGTTTGAACCTCGAGTCAACACGCTCT	3120	
Qy	3121	AGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTTGGAGGATGACCGG	3180	
Db	3121	AGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTTGGAGGATGACCGG	3180	
Qy	3181	TTTGGCAGCAGCTCTACTAGTACATTTTCTCTTGATATCAATGAAGATACAGAGCCAACA	3240	
Db	3181	TTTGGCAGCAGCTCTACTAGTACATTTTCTCTTGATATCAATGAAGATACAGAGCCAACA	3240	
Qy	3241	TTTTATGACCGATCTGGACCCATAAAGGATAAAAAATTCATTCCCTTTCTTTGCTCTTAGT	3300	
Db	3241	TTTTATGACCGATCTGGACCCATAAAGGATAAAAAATTCATTCCCTTTCTTTGCTCTTAGT	3300	
Qy	3301	AAACTGTGGAAGATCGTATCTTTAAATTCGCTTACTTTGCTTAACAAAAACAATCGTAGT	3360	
Db	3301	AAACTGTGGAAGATCGTATCTTTAAATTCGCTTACTTTGCTTAACAAAAACAATCGTAGT	3360	
Qy	3361	AGCAGTATCCAAAAGGAGGAAATTTATCATCTGAAGTAAAGACAGAAACAGGCGAATC	3420	
Db	3361	AGCAGTATCCAAAAGGAGGAAATTTATCATCTGAAGTAAAGACAGAAACAGGCGAATC	3420	
Qy	3421	AGAACCTTACGGAGCCCGAGTGTGATTTTAAATCATAGTGTGATGATTTTACCCCATATCC	3480	
Db	3421	AGAACCTTACGGAGCCCGAGTGTGATTTTAAATCATAGTGTGATGATTTTACCCCATATCC	3480	
Qy	3481	ACTGTACAGAAAAATTCATAATTAGAGACTTCATTTATGGGAAATGAAGACATTTGAAGAC	3540	
Db	3481	ACTGTACAGAAAAATTCATAATTAGAGACTTCATTTATGGGAAATGAAGACATTTGAAGAC	3540	
Qy	3541	ACTGGTAGTACACCAAGCATTTGAGAAATATGACTTTAAATTTCAACCAAGAAATTTTGGTACA	3600	
Db	3541	ACTGGTAGTACACCAAGCATTTGAGAAATATGACTTTAAATTTCAACCAAGAAATTTTGGTACA	3600	
Qy	3601	GAGAACTACAGAGAAATACCAAGCCGAGGAGGTAGTAGTGAAGAGTTTCAACGAGCTCA	3660	
Db	3601	GAGAACTACAGAGAAATACCAAGCCGAGGAGGTAGTAGTGAAGAGTTTCAACGAGCTCA	3660	
Qy	3661	CATATGAAGATAGGTAGCCAAAGTTTCAATACAGACACTACAAACAAGTGCATAAGTTCA	3720	
Db	3661	CATATGAAGATAGGTAGCCAAAGTTTCAATACAGACACTACAAACAAGTGCATAAGTTCA	3720	
Qy	3721	ATGAGCTCAAGTCTTCAACGAGACAGTAGTAGGTGTAGATGCTTCAACTATGGAACAGAC	3780	

Db 3721 ATGAGCTCAAGTCTCTTACGAGAGACAGTAGGTGTAGATGCTACAACTATGGACACAGAC 3780
Qy TGTGGAGAGTACGATCTGTGGTAAGTACTAAACTATTAAAGCAAGCCACTATTTCACG 3840
Db TGTGGAGAGTACGATCTGTGGTAAGTACTAAACTATTAAAGCAAGCCACTATTTCACG 3840
Qy CCACAGTCTAACCATCTGTCTCTCCAAATCAAATTCGGTGTCCCTGGTGCCTCCAGGT 3900
Db CCACAGTCTAACCATCTGTCTCTCCAAATCAAATTCGGTGTCCCTGGTGCCTCCAGGT 3900
Qy TCTTCTCATAGCTTCTAGAGAGCACAGTCCCTTAAAGCACCTCTATTTGCTACAAAT 3960
Db TCTTCTCATAGCTTCTAGAGAGCACAGTCCCTTAAAGCACCTCTATTTGCTACAAAT 3960
Qy AAAAGTCTACAGATTGTAACCTTTAGTTTACAGAGTTCTAGAGATGCTTTTGGCTATGCT 4020
Db AAAAGTCTACAGATTGTAACCTTTAGTTTACAGAGTTCTAGAGATGCTTTTGGCTATGCT 4020
Qy ACATGAAAAGACTACAGCAACAAAGAAATGCATCCCTTATCTCACTCTGAAGCTTTG 4080
Db ACATGAAAAGACTACAGCAACAAAGAAATGCATCCCTTATCTCACTCTGAAGCTTTG 4080
Qy GCATCTCCAGCAAAAGATGTCTATTTACTGTATACCATCACCATGAAGGCCAACAGTTT 4140
Db GCATCTCCAGCAAAAGATGTCTATTTACTGTATACCATCACCATGAAGGCCAACAGTTT 4140
Qy GAGTCCAGATTAACACCAAGCAGGTTCAAGAAAGCCTTAAGTTATGCATCATTTAGATAAA 4200
Db GAGTCCAGATTAACACCAAGCAGGTTCAAGAAAGCCTTAAGTTATGCATCATTTAGATAAA 4200
Qy GAAGATTTATTGAGTCTCTAATTAATCAAAATACCCCTGCAACGATCTTCTCAGTCCGGTCC 4260
Db GAAGATTTATTGAGTCTCTAATTAATCAAAATACCCCTGCAACGATCTTCTCAGTCCGGTCC 4260
Qy ATGGTGTCCAGTCCGCACATATGGGGGTTTCAGATGATTAATGATTTGATGATTTGATGAT 4320
Db ATGGTGTCCAGTCCGCACATATGGGGGTTTCAGATGATTAATGATTTGATGATTTGATGAT 4320
Qy GATATAAATGATATATTTCCAGGTAAAGGATATTCCTATTTTTCAGACAAATAACATACCA 4380
Db GATATAAATGATATATTTCCAGGTAAAGGATATTCCTATTTTTCAGACAAATAACATACCA 4380
Qy CCACATGATGATCAGAGTGCAGAGCATTTGCCCATGATGCAGAGGTTCTTCATCTGGA 4440
Db CCACATGATGATCAGAGTGCAGAGCATTTGCCCATGATGCAGAGGTTCTTCATCTGGA 4440
Qy ACTGGAGGTTCTGTAAAAAATCTTTTCACTGTCTACGACAGCATGATGCTTACGGAA 4500
Db ACTGGAGGTTCTGTAAAAAATCTTTTCACTGTCTACGACAGCATGATGCTTACGGAA 4500
Qy ATAAATCAATTCATCCATTCAGATGCTCTCTGTTTTTAAAGTACAGACACTGGA 4560
Db ATAAATCAATTCATCCATTCAGATGCTCTCTGTTTTTAAAGTACAGACACTGGA 4560
Qy CTACAGGAACATACAGATGATACTGCTTTATTTGTTGCTGTATTTGAAATTTCTGGTTTC 4620
Db CTACAGGAACATACAGATGATACTGCTTTATTTGTTGCTGTATTTGAAATTTCTGGTTTC 4620
Qy CAGCCAGCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTCAAGATATTCATAT 4680
Db CAGCCAGCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTCAAGATATTCATAT 4680
Qy TCTGATTTGTTGAGCAGACTATCCATAATCTTTTAAAGTGTCTCCCTCAAGTTTTCG 4740
Db TCTGATTTGTTGAGCAGACTATCCATAATCTTTTAAAGTGTCTCCCTCAAGTTTTCG 4740
Qy GGGATTTCTGGATGCACTAGTGGGTGCTCAGAGAGGCTCAGCTAGCAGCACCACCAAGC 4800
Db GGGATTTCTGGATGCACTAGTGGGTGCTCAGAGAGGCTCAGCTAGCAGCACCACCAAGC 4800
Qy ACAGATTTCTTACTAGTGTAAACCAATTCAGATGATACCAATGTCGGTATCTC 4860
Db ACAGATTTCTTACTAGTGTAAACCAATTCAGATGATACCAATGTCGGTATCTC 4860

Qy 4861 CTTTCGCAAGAGAGTTCTTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAACTAAATGT 4920
Db CTTTCGCAAGAGAGTTCTTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAACTAAATGT 4920
Qy CATGAGACTGGGCTTTTAAACAATTAAAGGAGAGATATCCTCAACAATTTTGATGACATATGC 4980
Db CATGAGACTGGGCTTTTAAACAATTAAAGGAGAGATATCCTCAACAATTTTGATGACATATGC 4980
Qy CTTTACTCTGAGGTTTCCCATTTTGTCTGCACACTGCACATTCAGACTTCGGTTCGGAGG 5040
Db CTTTACTCTGAGGTTTCCCATTTTGTCTGCACACTGCACATTCAGACTTCGGTTCGGAGG 5040
Qy TTCAATACAGAATTTATTTCAAGATGTACAGTTTCTCAAAATGCAAGAGCAGAGGCT 5100
Db TTCAATACAGAATTTATTTCAAGATGTACAGTTTCTCAAAATGCAAGAGCAGAGGCT 5100
Qy GTGTTGGCAACACCACCAAGCAACCTATAGTTGTATACATCTGCTGAATCCTGACCTCAT 5160
Db GTGTTGGCAACACCACCAAGCAACCTATAGTTGTATACATCTGCTGAATCCTGACCTCAT 5160
Qy ATTTATGATGGATATAGATACATCTATATATATTTTGTGATTTTCTTAAAGGCC 5220
Db ATTTATGATGGATATAGATACATCTATATATATTTTGTGATTTTCTTAAAGGCC 5220
Qy TCAGAAAATACGACTGACTAGGAGCAAGACAGAGATATCTTCTGTACACTGTTCCGCA 5280
Db TCAGAAAATACGACTGACTAGGAGCAAGACAGAGATATCTTCTGTACACTGTTCCGCA 5280
Qy GTTACTGGTACATGAACAGTTCGAACTGCTGAACTTTCTTAACCAAAACAACCTTCTTCTC 5340
Db GTTACTGGTACATGAACAGTTCGAACTGCTGAACTTTCTTAACCAAAACAACCTTCTTCTC 5340
Qy TCTTTGTTGAGCCTTTTGGAGGGTTTCATGATTCATTACCACAGTTTAAAGAGTTTCAGT 5400
Db TCTTTGTTGAGCCTTTTGGAGGGTTTCATGATTCATTACCACAGTTTAAAGAGTTTCAGT 5400
Qy TACCAATTTGATGCAAGAGCAAGCACTGAATACCTACATAGGTTTCTATTTTCTTTTCA 5460
Db TACCAATTTGATGCAAGAGCAAGCACTGAATACCTACATAGGTTTCTATTTTCTTTTCA 5460
Qy TTTTAAAGCGTAAATGACAGTGGAACTAATGGGATATGCAAGCACCCTTTCACAAGTT 5520
Db TTTTAAAGCGTAAATGACAGTGGAACTAATGGGATATGCAAGCACCCTTTCACAAGTT 5520
Qy ATTTCTGAATGATTTTAGGGTAAATAAATACAGATGCTTGTATGTAACTAACTTGTGGA 5580
Db ATTTCTGAATGATTTTAGGGTAAATAAATACAGATGCTTGTATGTAACTAACTTGTGGA 5580
Qy AAGCAGGAATCAGTCTCTTAAGGCTGCATCTTATTCACAATGGGTTGTGCTATAC 5640
Db AAGCAGGAATCAGTCTCTTAAGGCTGCATCTTATTCACAATGGGTTGTGCTATAC 5640
Qy TGGCTGGTATTAGAGAGGGAAC 5662
Db TGGCTGGTATTAGAGAGGGAAC 5662

RESULT 2
ID ADRA4435

XX ADRA4435 standard; cDNA; 5127 BP.

AC ADRA4435;

DT 18-NOV-2004 (first entry)

DE Human p200 encoding cDNA SEQ ID NO:2.

KW mTOR-associated protein; mTOR-AP; cytostatic; antidiabetic; gene therapy;

KW mTOR-AP modulation; cancer; tumour; diabetes; human; p200; gene; ss.

OS Homo sapiens.

XX

PH	Key	Location/Qualifiers	
FT	CDS	1..5127	
FT		/*tag= a	
FT		/product= "p200"	
XX		WO2004074448-A2.	
PN		02-SEP-2004.	
PD		18-FEB-2004; 2004WO-US004821.	
PF		18-FEB-2003; 2003US-0448035P.	
PP		(WHED) WHITEHEAD INST BIOMEDICAL RES.	
PR		Sabatini DM, Kim D, Sarbaessov D;	
XX		WPI; 2004-635565/61.	
XX		N-FSDB; ADR44435.	
PA		New mTOR-associated protein (mTOR-AP) polypeptide, useful for treating or	
XX		preventing a disorder that is responsive to mTOR-AP modulation, e.g.	
XX		cancer or diabetes.	
PT		Claim 1; SEQ ID NO 2; 91pp; English.	
PT		The present invention describes an isolated mTOR-associated protein (mTOR	
XX		-AP) polypeptide comprising a sequence that is 70% identical to the 1708	
CC		amino acid sequence of SEQ ID NO:3 (ADR44435, P1) or 70% identical to the	
CC		327 amino acid sequence of SEQ ID NO:6 (ADR44439, P2), or that is encoded	
CC		by a nucleic acid that hybridizes under high stringency conditions to a	
CC		complement sequence of the 951271 nucleic acid sequence of SEQ ID NO:2	
CC		(ADR44435, S1) or the 981 nucleic acid sequence of SEQ ID NO:5 (ADR44438,	
CC		S2). Also described: (1) an isolated antibody, or its fragment, which is	
CC		specifically immunoreactive with an epitope of an amino acid sequence of	
CC		P1 or P2; (2) the isolated nucleic acid encoding the polypeptide, or its	
CC		complement; (3) a vector comprising a nucleic acid sequence encoding an	
CC		mTOR-AP polypeptide; (4) an isolated host cell comprising the nucleic	
CC		acid; (5) a method of producing an mTOR-AP polypeptide; (6) a method for	
CC		detecting the presence of an mTOR-AP polypeptide; (7) a kit for detecting	
CC		an mTOR-AP polypeptide comprising the antibody and a detectable label for	
CC		detecting the antibody; (8) a method for detecting the presence of the	
CC		nucleic acid in a sample; (9) a kit comprising the probe or primer	
CC		nucleic acid and instructions for use; (10) an isolated, purified or	
CC		recombinant complex comprising an mTOR polypeptide and an mTOR-AP; (11) a	
CC		method for identifying a compound which modulates activity of an mTOR-AP	
CC		polypeptide; (12) a method of inhibiting aberrant activity of an mTOR-AP-	
CC		expressing cell; (13) a method of treating or preventing a disorder that	
CC		is responsive to mTOR-AP modulation, in a subject; and (14) a transgenic	
CC		mouse having germline and somatic cells comprising a chromosomally	
CC		incorporated transgene that disrupts the genomic mTOR-AP gene and	
CC		inhibits expression of the gene, where the disruption comprises insertion	
CC		of a selectable marker sequence, resulting in the transgenic mouse	
CC		exhibiting increased susceptibility to the formation of tumours as	
CC		compared to the wild type mouse. mTOR-AP sequences cytostatic and	
CC		antidiabetic activities, and can be used in gene therapy. The	
CC		polypeptides, polynucleotides, compounds and methods are useful for	
CC		treating or preventing a disorder that is responsive to mTOR-AP	
CC		modulation, e.g. cancer or diabetes. The compound, which modulates an	
CC		mTOR-AP activity or expression is useful for the manufacture of a	
CC		medicament for treating a disease affected by aberrant mTOR-AP activity	
CC		or expression. The present sequence encodes human p200, which is an mTOR-	
CC		AP used in the exemplification of the present invention.	
XX		Sequence 5127 BP; 1626 A; 1028 C; 1049 G; 1424 T; 0 U; 0 Other;	
SQ		Query Match 90.6%; Score 5127; DB 13; Length 5127;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy		28 ATGCGCGGATCGCGCGCGCGCTCTCTGAGAACCTCCGAGTACGAGGCGGAATGAC 87	
Db		1 ATGCGCGGATCGCGCGCGCGCTCTCTGAGAACCTCCGAGTACGAGGCGGAATGAC 60	

Qy	88	AGCGCGAGGAGAACGTCGCCCTCGATCTGACCCGAGAACCTTCTGTATAAATTAAGAGAG	147
Db	61	AGCGCGAGGAGAACGTCGCCCTCGATCTGACCCGAGAACCTTCTGTATAAATTAAGAGAG	120
Qy	148	ATTCTCCAAATATGGCCAGATTCGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG	207
Db	121	ATTCTCCAAATATGGCCAGATTCGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG	180
Qy	208	AATAACTTTTACTAAGCTTTCTTTGTATATTGGCCACAGTGAAGAAAAGCTGGCTTTTAC	267
Db	181	ATAAATTTTACTAAGCTTTCTTTGTATATTGGCCACAGTGAAGAAAAGCTGGCTTTTAC	240
Qy	268	TATGAGGATATCAATAATTTGTTTTCGGTTAGCTTTTATTAATGAAGCAAAAGAGTGGCA	327
Db	241	TATGAGGATATCAATAATTTGTTTTCGGTTAGCTTTTATTAATGAAGCAAAAGAGTGGCA	300
Qy	328	GCAGCAGGCTACGAGGCTTCGATATCTCATCAAGACTCCAGTATTCTCCAGAAGTG	387
Db	301	GCAGCAGGCTACGAGGCTTCGATATCTCATCAAGACTCCAGTATTCTCCAGAAGTG	360
Qy	388	CTAAATTTGAAGTGGACTATTTAATAGTAGTGGTGCATTGACATACACAGCAACGAG	447
Db	361	CTAAATTTGAAGTGGACTATTTAATAGTAGTGGTGCATTGACATACACAGCAACGAG	420
Qy	448	GTAGAGAGGACACAAAGCACTTCGATTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG	507
Db	421	GTAGAGAGGACACAAAGCACTTCGATTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG	480
Qy	508	TTTCCTAGTCTGTGACCAACTCAATTAATGTCAGTGGAAATGATGGACTTCAAGAAAGA	567
Db	481	TTTCCTAGTCTGTGACCAACTCAATTAATGTCAGTGGAAATGATGGACTTCAAGAAAGA	540
Qy	568	GACAGATGTCGCGAGCATGCTGATTCGATTCGTGAACTAGGACTTCAGAAATCCAGAG	627
Db	541	GACAGATGTCGCGAGCATGCTGATTCGATTCGTGAACTAGGACTTCAGAAATCCAGAG	600
Qy	628	GTGTGCGCCCTTCGAGGAGGACTAAACCAATATTTGAAATAATGATTTGATTTGCAATTA	687
Db	601	GTGTGCGCCCTTCGAGGAGGACTAAACCAATATTTGAAATAATGATTTGATTTGCAATTA	660
Qy	688	AGTCGAATAAATGAGGCGCTTAATTTACTAATTTTTCACCTCTTAAATCATCCAAAGACT	747
Db	661	AGTCGAATAAATGAGGCGCTTAATTTACTAATTTTTCACCTCTTAAATCATCCAAAGACT	720
Qy	748	CGGCAATGTCGCGAGCTGATAGTAATTTAGAGAGTAATTTAGCACCTTACTGATTTT	807
Db	721	CGGCAATGTCGCGAGCTGATAGTAATTTAGAGAGTAATTTAGCACCTTACTGATTTT	780
Qy	808	CACCTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAGAAAGACAGAGAACGCA	867
Db	781	CACCTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAGAAAGACAGAGAACGCA	840
Qy	868	TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGATCATGGGAGGATTTATTAAAT	927
Db	841	TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGATCATGGGAGGATTTATTAAAT	900
Qy	928	TTATGTAACCTGGAAATTTCTGGGATCCAGTCTCTAATAGGAGTACTTTGCAATACCAAT	987
Db	901	TTATGTAACCTGGAAATTTCTGGGATCCAGTCTCTAATAGGAGTACTTTGCAATACCAAT	960
Qy	988	ATGGAATAAGGCGAGGCTTACTTTGAAGTCTTTATGATATATTTTCGTTCTCTCTACCT	1047
Db	961	ATGGAATAAGGCGAGGCTTACTTTGAAGTCTTTATGATATATTTTCGTTCTCTCTACCT	1020
Qy	1048	GTTGTGACTGAGGATTCATAGAACCACTACTCAGTGTAGATCCAGGAGGTTCCAAAGAC	1107
Db	1021	GTTGTGACTGAGGATTCATAGAACCACTACTCAGTGTAGATCCAGGAGGTTCCAAAGAC	1080
Qy	1108	AGTTGGAGGCTTTCAGATGCTTTGTGGCAGCTGAGGCAAAAGCTATTTCTCTCATCGT	1167
Db	1081	AGTTGGAGGCTTTCAGATGCTTTGTGGCAGCTGAGGCAAAAGCTATTTCTCTCATCGT	1140

1168 GCAGATCCAGGCGAGACCTCATGGATAAATATTTGGCACTGATACCTCTCTGATTTATTT 1227
1141 GCCAGATCCAGGCGAGACCTCATGGATTAATATTTGGCACTGATACCTCTCTGATTTATTT 1200
1228 CGTAATGGAATTTAGAGGCTCTAGTTGAAGTGAATAACAAACAGTGAATCATATCTCA 1287
1201 CGTAATGGAATTTAGAGGCTCTAGTTGAAGTGAATAACAAACAGTGAATCATATCTCA 1260
1288 GTTAGAGCTACCATCTTTTAGAGAGCTTTTACATATGCAACACAACTTTCTCTCAT 1347
1261 GTTAGAGCTACCATCTTTTAGAGAGCTTTTACATATGCAACAACTTTCTCTCAT 1320
1348 TCATATAGCCATCATTTTACACTGCTTGCACAACTTAATGAATGCTGCTCATCTTTGAT 1407
1321 TCATATAGCCATCATTTTACACTGCTTGCACAACTTAATGAATGCTGCTCATCTTTGAT 1380
1408 ATCCCAAGAAAGAGATGCGAGCGAGTGCAGCCTTGAATCTGTTTAAACAGCTTCCAT 1467
1381 ATCCCAAGAAAGAGATGCGAGCGAGTGCAGCCTTGAATCTGTTTAAACAGCTTCCAT 1440
1468 GAAATGAAGAAAGAGATGCGAGCGAGTGCAGCCTTGAATCTGTTTAAACAGCTTCCAT 1527
1441 GAAATGAAGAAAGAGATGCGAGCGAGTGCAGCCTTGAATCTGTTTAAACAGCTTCCAT 1500
1528 GCAATTCGAACACACACAGAAACGGGATCAGTATCTCCGAGTTCAGAAAGATATATTTATC 1587
1501 GCAATTCGAACACACACAGAAACGGGATCAGTATCTCCGAGTTCAGAAAGATATATTTATC 1560
1588 CTTAAGGATACAGAGGAGCTCTTTTAAATTAATTAACCTTAGAGATAGCCAGTCTTCAACAT 1647
1561 CTTAAGGATACAGAGGAGCTCTTTTAAATTAATTAACCTTAGAGATAGCCAGTCTTCAACAT 1620
1648 AAGAGAACTTCAATGGAATTTGGAATCTTATAGGACCACTTCTTAAGTGGCCAAATGTA 1707
1621 AAGAGAACTTCAATGGAATTTGGAATCTTATAGGACCACTTCTTAAGTGGCCAAATGTA 1680
1708 AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT 1767
1681 AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT 1740
1768 TACAAGCCGAGAGTAAATTAATGCAACCTTGGATCTGGATTTTCCAAAGGCCAAACAG 1827
1741 TACAAGCCGAGAGTAAATTAATGCAACCTTGGATCTGGATTTTCCAAAGGCCAAACAG 1800
1828 CTCACGGTTGTAGGTTGCCAGTTTACAGAAATTTCTTGAATCTGGAAGAGATGGCAA 1887
1801 CTCACGGTTGTAGGTTGCCAGTTTACAGAAATTTCTTGAATCTGGAAGAGATGGCAA 1860
1888 GGCTACTTGAAGATCTAGTAAAGGATATTTGTTCACTGGCTCAATGCTTTCATCTGGAATG 1947
1861 GGCTACTTGAAGATCTAGTAAAGGATATTTGTTCACTGGCTCAATGCTTTCATCTGGAATG 1920
1948 AAACCCGAAAGAGTCTTCAAAATATGTTTATTTAGCACCCCTTAGTCAACACTACTTTT 2007
1921 AAACCCGAAAGAGTCTTCAAAATATGTTTATTTAGCACCCCTTAGTCAACACTACTTTT 1980
2008 TTATTTATTTGGAACACTTTCTTGCCACCTCTAGGAGTTAAATGCTGGAAGATGGCAGT 2067
1981 TTATTTATTTGGAACACTTTCTTGCCACCTCTAGGAGTTAAATGCTGGAAGATGGCAGT 2040
2068 GTATTTCACTGTCTCTTAATCTTTGCTCTTGAATAACCAAGATCACTTTGCTAAAACCTT 2127
2041 GTATTTCACTGTCTCTTAATCTTTGCTCTTGAATAACCAAGATCACTTTGCTAAAACCTT 2100
2128 ACTGTTCTTAGCTTGAACATATAGAGAGATGGAATTTGGCTAGAGTCACTTTTCCAAATTT 2187
2101 ACTGTTCTTAGCTTGAACATATAGAGAGATGGAATTTGGCTAGAGTCACTTTTCCAAATTT 2160
2188 TTAACTGCAGCTACTGATCCCTGCAGCTCTATGCAACAAACATTTTAAGGGTATTTATG 2247
2161 TTAACTGCAGCTACTGATCCCTGCAGCTCTATGCAACAAACATTTTAAGGGTATTTATG 2220
2248 AGAGCTAATGTTGAAATTTTAAATTAATGGGAATTTGAGTTGTTAGTGACCCAGCTACAT 2307

2221 AGAGCTAATGTTGAATCTTTTAATTAATTTGGGAATTTGAGTTGTTAGTGAACCCAGCTACAT 2280
2308 GATAAAAAACAAACGATTTTCTCTGAAGCTCTTTGATATATCTCGATGAAGCATGTGAAGAC 2367
2281 GATAAAAAACAAACGATTTTCTCTGAAGCTCTTTGATATATCTCGATGAAGCATGTGAAGAC 2340
2368 AAGGCCAATCTTCATGCTCTCTCAATTCAGATGAACACAGCGTTATCCACCTTTGGAGACAAG 2427
2341 AAGGCCAATCTTCATGCTCTCTCAATTCAGATGAACACAGCGTTATCCACCTTTGGAGACAAG 2400
2428 GGTTCCTCTCTCTGCTGAGATTTCTCTCCATTTCCAAAAAGGATTTTCTCTATCTGTAATGAA 2487
2401 GGTTCCTCTCTCTGCTGAGATTTCTCTCCATTTCCAAAAAGGATTTTCTCTATCTGTAATGAA 2460
2488 AGAGGTTATGTAGCAAAACAAATTTGGAAAAAGTGGCAACAGGAATACAACTCCAAATATGTT 2547
2461 AGAGGTTATGTAGCAAAACAAATTTGGAAAAAGTGGCAACAGGAATACAACTCCAAATATGTT 2520
2548 GACTTGATTTGAGGAACAACTCAATGAAGCACTTACTACTTACCGGAAGCTGTTGATGTT 2607
2521 GACTTGATTTGAGGAACAACTCAATGAAGCACTTACTACTTACCGGAAGCTGTTGATGTT 2580
2608 GATAACTATGTTTCGTCGGAGTAACCAAGATTTACAGCGTCTCACGTCTACCTGCCCTATA 2667
2581 GATAACTATGTTTCGTCGGAGTAACCAAGATTTACAGCGTCTCACGTCTACCTGCCCTATA 2640
2668 CACCTTTATFGACAACCTAGTACACCAATAAAACAGGCTGCCATTTGTTGGAAGTACAGAAT 2727
2641 CACCTTTATFGACAACCTAGTACACCAATAAAACAGGCTGCCATTTGTTGGAAGTACAGAAT 2700
2728 ATTATTTACAGAACTCTGTCGTAATGTTTGTACACAGATTTTGGATAAGTGGGAAGAAATTT 2787
2701 ATTATTTACAGAACTCTGTCGTAATGTTTGTACACAGATTTTGGATAAGTGGGAAGAAATTT 2760
2788 AAAAAAATGAAAGCATCTTTTGGGCTTTGGGAATATCGGCTCATCAAAATTTGGGCTCTC 2847
2761 AAAAAAATGAAAGCATCTTTTGGGCTTTGGGAATATCGGCTCATCAAAATTTGGGCTCTC 2820
2848 AATTTGCTACAGGAAGAAAACTGATTCAGATATACATAAATCTTGCAAAAACAGTGTGAA 2907
2821 AATTTGCTACAGGAAGAAAACTGATTCAGATATACATAAATCTTGCAAAAACAGTGTGAA 2880
2908 GTTCTTTTCCATCAGAGGGACCTGCTATATGTAATTTGGGCTCATAGCTTAAACCAACAA 2967
2881 GTTCTTTTCCATCAGAGGGACCTGCTATATGTAATTTGGGCTCATAGCTTAAACCAACAA 2940
2968 GGCTGTGATATTTCTAAAATGTCAAACTGGGATGCTGTGAGGCATAGTGCAGCAACATCTG 3027
2941 GGCTGTGATATTTCTAAAATGTCAAACTGGGATGCTGTGAGGCATAGTGCAGCAACATCTG 3000
3028 TGGCCAGTGGTTCCAGATGATGTGAAACAACTCTGTGAATGAACCTTTTCACTATCCCAAGC 3087
3001 TGGCCAGTGGTTCCAGATGATGTGAAACAACTCTGTGAATGAACCTTTTCACTATCCCAAGC 3060
3088 ACTCTAAGTTTGAACTCGGAGTCAACAGCTCTAGACATAATAGTGAAGTCAATCTG 3147
3061 ACTCTAAGTTTGAATCTCGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTG 3120
3148 CCATCGAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3207
3121 CCATCGAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
3208 TTCCTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCCAATAAG 3267
3181 TTCCTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCCAATAAG 3240
3268 GATAAAAAATCAATCT 3327
3241 GATAAAAAATCAATCT 3300
3328 TCGCTTACTTTTGCCTTAACAAAAACATCTGATAGTACAGTGTGATTCAAAAAGGGGAAATTA 3387

Db 3301 TCGCTTACTTGGCTTAA CAAAAAACA TCGTAGTAGCAGTGATCCAAAAGGAGGAAAAATTA 3360
Qy 3388 TCATCTGAAGTAAGACAAGCAACAGGCGGAATCAGAAACACTTACGGAGCCCAAGTGTGAT 3447
Db 3361 TCATCTGAAGTAAGACAAGCAACAGGCGGAATCAGAAACACTTACGGAGCCCAAGTGTGAT 3420
Qy 3448 TTTAATCATAGTAGATTTTACACCCCATATCCACTGTACAGAAACAAATTAACAATTAGAG 3507
Db 3421 TTTAATCATAGTAGATTTTACACCCCATATCCACTGTACAGAAACAAATTAACAATTAGAG 3480
Qy 3508 ACTTCATTTATGGGGATAAGACACATTTGAAGACACTGTGTAGTACACCAAGCATTTGGAGAA 3567
Db 3481 ACTTCATTTATGGGGATAAGACACATTTGAAGACACTGTGTAGTACACCAAGCATTTGGAGAA 3540
Qy 3568 AATGACTTTAAATTCACCAAGAAATTTTGGTACAGAGAAATCAGAGAAATATACAGCCGA 3627
Db 3541 AATGACTTTAAATTCACCAAGAAATTTTGGTACAGAGAAATCAGAGAAATATACAGCCGA 3600
Qy 3628 GAGAGGTTAGTAGTAAAGTTTCAACGAGCTTCAATATGAAGATACGTAGTCAAGTTTC 3687
Db 3601 GAGAGGTTAGTAGTAAAGTTTCAACGAGCTTCAATATGAAGATACGTAGTCAAGTTTC 3660
Qy 3688 AATACAGACACTACCAAGTGGCATAGTTTCAATGAGCTCAAGTCTTCCAGAGAGACA 3747
Db 3661 AATACAGACACTACCAAGTGGCATAGTTTCAATGAGCTCAAGTCTTCCAGAGAGACA 3720
Qy 3748 GTAGGTGTAGATGCTACAACTATGGACACAGACTGTGGAAGCATGAGTACTGTGTAAGT 3807
Db 3721 GTAGGTGTAGATGCTACAACTATGGACACAGACTGTGGAAGCATGAGTACTGTGTAAGT 3780
Qy 3808 ACTAAAACTATTAAGACAAGCCACTATTTGACGCCACAGTCTAACCTATCTCTCTCC 3867
Db 3781 ACTAAAACTATTAAGACAAGCCACTATTTGACGCCACAGTCTAACCTATCTCTCTCC 3840
Qy 3868 AAATCAAAATTCGGTCCCTGGTCCCTCCAGTTCTTCTCATAGCTTCTCTAGAGAGCA 3927
Db 3841 AAATCAAAATTCGGTCCCTGGTCCCTCCAGTTCTTCTCATAGCTTCTCTAGAGAGCA 3900
Qy 3928 CAGTCCCTTAAAGCACCTCTATTTGCTACAAATTAAGAGTCTAGCAGAGTTGTAACCTTTAGT 3987
Db 3901 CAGTCCCTTAAAGCACCTCTATTTGCTACAAATTAAGAGTCTAGCAGAGTTGTAACCTTTAGT 3960
Qy 3988 TACAAGAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA 4047
Db 3961 TACAAGAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA 4020
Qy 4048 ATGCATCCATCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4107
Db 4021 ATGCATCCATCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4080
Qy 4108 ACTGATACCATCACCATGAAGGCCAACAGATTTTTCAGTCCAGATTAAACCAAGCAGGTTTC 4167
Db 4081 ACTGATACCATCACCATGAAGGCCAACAGATTTTTCAGTCCAGATTAAACCAAGCAGGTTTC 4140
Qy 4168 ATGAAGCCCTTAAGTTATGATCATATTAGATAAAGAGATTTATTGAGTCTCTATTAAATCAA 4227
Db 4141 ATGAAGCCCTTAAGTTATGATCATATTAGATAAAGAGATTTATTGAGTCTCTATTAAATCAA 4200
Qy 4228 AATACCCCTGCAACGATCTTCTCAGTGGGTCCATGGTGTCCAGTGCACATATGGGGGT 4287
Db 4201 AATACCCCTGCAACGATCTTCTCAGTGGGTCCATGGTGTCCAGTGCACATATGGGGGT 4260
Qy 4288 TCAGATGATTACATTTGCTTGTCTCTCCGGTGGATATAATGATATATTTCCAGGTAAAG 4347
Db 4261 TCAGATGATTACATTTGCTTGTCTCTCCGGTGGATATAATGATATATTTCCAGGTAAAG 4320
Qy 4348 GATATTCCTTATTTTCAGACAAAAAATACACCATGATGATCGAGGTGCAAGAGCA 4407
Db 4321 GATATTCCTTATTTTCAGACAAAAAATACACCATGATGATCGAGGTGCAAGAGCA 4380
Qy 4408 TTTGCCCATGATGAGGAGGTCTTCCATCTGGAACTCGGAGGTCTTGTAAAAAATCTTTTT 4467
Db 4381 TTTGCCCATGATGAGGAGGTCTTCCATCTGGAACTCGGAGGTCTTGTAAAAAATCTTTTT 4440

Qy 4468 CACTTGTCTACGACAGCAGATGAGTCTTACGGAAAAATTAATTCATTCATTCAGATGCC 4527
Db 4441 CACTTGTCTACGACAGCAGATGAGTCTTACGGAAAAATTAATTCATTCATTCAGATGCC 4500
Qy 4528 TCTCTGTTTTTAGAAAAGTACAGAAAGACACTGGACTACAGGAACATACAGATGATAACTGC 4587
Db 4501 TCTCTGTTTTTAGAAAAGTACAGAAAGACACTGGACTACAGGAACATACAGATGATAACTGC 4560
Qy 4588 CTTTATTGTCTGTATTGAAATTTCTGGGTTTCCAGCCCAAGCAACTGAGTGCATA 4647
Db 4561 CTTTATTGTCTGTATTGAAATTTCTGGGTTTCCAGCCCAAGCAACTGAGTGCATA 4620
Qy 4648 TGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATGGTGTGAGCAGACTATCCAT 4707
Db 4621 TGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATGGTGTGAGCAGACTATCCAT 4680
Qy 4708 AATCCTTTAGAAAGTGTCCCTCTAAGTTTTTCGGGATTTTCTGATGACGTGATGGGTG 4767
Db 4681 AATCCTTTAGAAAGTGTTCCTCTAAGTTTTTCGGGATTTTCTGATGACGTGATGGGTG 4740
Qy 4768 TCTCAAGAAGGCTCAGCTAGCAGCACCACCAAGACACAGAAATTTGTTACTAGTGTAAACA 4827
Db 4741 TCTCAAGAAGGCTCAGCTAGCAGCACCACCAAGACACAGAAATTTGTTACTAGTGTAAACA 4800
Qy 4828 ATTCCAGATGATACACCAATGTGCCGTATACTCTCTCGCAAGAAAGTTCTAAGATTAGTC 4887
Db 4801 ATTCCAGATGATACACCAATGTGCCGTATACTCTCTCGCAAGAAAGTTCTAAGATTAGTC 4860
Qy 4888 ATTAATTTGAGTAGTTCAGTTTTCAACTAAATGTATGAGACTGGGCTTTTAAACAATTAAAG 4947
Db 4861 ATTAATTTGAGTAGTTCAGTTTTCAACTAAATGTATGAGACTGGGCTTTTAAACAATTAAAG 4920
Qy 4948 GAGAAGTATCCTCAAAATTTGATGACATATGCTTTACTCTGAGGTTTCCCATTTGCTG 5007
Db 4921 GAGAAGTATCCTCAAAATTTGATGACATATGCTTTACTCTGAGGTTTCCCATTTGCTG 4980
Qy 5008 TCACACTGACATTCAGACTTTCCGTGTCCGAGGTTTCATACAGAAATTTATTTCAAGATGTA 5067
Db 4981 TCACACTGACATTCAGACTTTCCGTGTCCGAGGTTTCATACAGAAATTTATTTCAAGATGTA 5040
Qy 5068 CAGTTTCTACAAATGCAATGAAGAGCAGAGGCTGTGTGGCAACACCAACCAAGCAACCT 5127
Db 5041 CAGTTTCTACAAATGCAATGAAGAGCAGAGGCTGTGTGGCAACACCAACCAAGCAACCT 5100
Qy 5128 ATAGTTGATACATCTGCTGAATCCTGA 5154
Db 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127

RESULT 3

AD71191
ID ADE71191 standard; DNA; 8232 BP.
XX
AC ADE71191;
XX AC
DT 29-JAN-2004 (first entry)
XX
DE Novel human protein coding sequence #7.
XX human; novel protein; drug; gene; ds.
XX Homo sapiens.
XX OS
XX JP2002345493-A.
XX
XX 03-DEC-2002.
XX PD
XX 29-MAR-2001; 2002JP-00049046.
XX PF
XX 29-MAR-2001; 2001JP-00095524.
XX PR
XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.

[illegible]

QY 3227 ATACAGAGCCAACTTTTATGACCGATCTGGACCCATAAAGGATPAAAAATTCATTCCCTT 3286
DB 1921 ATACAGAGCCAACTTTTATGACCGATCTGGACCCATAAAGGATPAAAAATTCATTCCCTT 1980
QY 3287 TCTTTGCTTCTAGTAAACTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAACA 3346
DB 1981 TCTTTGCTTCTAGTAAACTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAACA 2040
QY 3347 AAAAACTCGTAGCAGTATCAAAAGGAGGAAATATCATCTGAAAGTAAAGACAA 3406
DB 2041 AAAAACTCGTAGCAGTATCAAAAGGAGGAAATATCATCTGAAAGTAAAGACAA 2100
QY 3407 GCAACGCGGAATCAGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGATGATT 3466
DB 2101 GCAACGCGGAATCAGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGATGATT 2160
QY 3467 TTACACCCATATCCACTGTACAGAAAACATTAAGTATGAGACTTCAATTTATGGGGAATA 3526
DB 2161 TTACACCCATATCCACTGTACAGAAAACATTAAGTATGAGACTTCAATTTATGGGGAATA 2220
QY 3527 AGCACTTTGAAGACACTGGTAGTACCAAGCATTTGGAGAAAATGACTTAAATTCACCA 3586
DB 2221 AGCACTTTGAAGACACTGGTAGTACCAAGCATTTGGAGAAAATGACTTAAATTCACCA 2280
QY 3587 AGAATTTTGTGTACAGAGAAATCAGAGAGAAATACAAGCCGAGAGAGGTTAGTAGTAAAA 3646
DB 2281 AGAATTTTGTGTACAGAGAAATCAGAGAGAAATACAAGCCGAGAGAGGTTAGTAGTAAAA 2340
QY 3647 GTTCAACGAGCTACATATGAAGTACGTAGCCAAAGTTTCAATACAGACACTAACACAA 3706
DB 2341 GTTCAACGAGCTACATATGAAGTACGTAGCCAAAGTTTCAATACAGACACTAACACAA 2400
QY 3707 GTGSCATAAGTTCAATCAGCTCAAGTCTTACAGAGACAGTAGGTAGTAGTACTACAA 3766
DB 2401 GTGSCATAAGTTCAATCAGCTCAAGTCTTACAGAGACAGTAGGTAGTAGTACTACAA 2460
QY 3767 CTATGGACACAGACTGTGGAGCAGTAGTACTGTGTAAAGTACTTAAACTATTAAAGACAA 3826
DB 2461 CTATGGACACAGACTGTGGAGCAGTAGTACTGTGTAAAGTACTTAAACTATTAAAGACAA 2520
QY 3827 GCCACTATTTGACGCCACAGCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCC 3886
DB 2521 GCCACTATTTGACGCCACAGCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCC 2580
QY 3887 TGGTGCTCCAGCTTCTCTCATACGCTTCTAGAGACACAGTCCCTTTAAAGCACCT 3946
DB 2581 TGGTGCTCCAGCTTCTCTCATACGCTTCTAGAGACACAGTCCCTTTAAAGCACCT 2640
QY 3947 CTATTTGCTACAAATTAAGTCTAGCAGATTGTAACTTTAGTTTACACAAAGTTCTAGAGATG 4006
DB 2641 CTATTTGCTACAAATTAAGTCTAGCAGATTGTAACTTTAGTTTACACAAAGTTCTAGAGATG 2700
QY 4007 CTTTTGCTATGCTACACTGAAAGACTACAGCAACAAAGATGCAATCCATCTTATCTC 4066
DB 2701 CTTTTGCTATGCTACACTGAAAGACTACAGCAACAAAGATGCAATCCATCTTATCTC 2760
QY 4067 ACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGA 4126
DB 2761 ACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGA 2820
QY 4127 AGGCCAACAGTTTTGAGTCAGATTTAACACCAAGCAGGTTTCATGAAGCCCTTAAAGTTATG 4186
DB 2821 AGGCCAACAGTTTTGAGTCAGATTTAACACCAAGCAGGTTTCATGAAGCCCTTAAAGTTATG 2880
QY 4187 CATCATPAGATAAAGAAATTTATTTAGTCTTATTAATCAAAATPACCCCTGCAACGATCTT 4246
DB 2881 CATCATPAGATAAAGAAATTTATTTAGTCTTATTAATCAAAATPACCCCTGCAACGATCTT 2940
QY 4247 CCTCAGTCCGCTCCATGGTCTCAGTGCCACATATGGGGTTTCAATGATTACATTGGTC 4306
DB 2941 CCTCAGTCCGCTCCATGGTCTCAGTGCCACATATGGGGTTTCAATGATTACATTGGTC 3000

QY 4307 TTGCTCTCCGCTGGATATAAATGATATATTCAGGTAAAGGATATTCCTTATTTTCAGA 4366
DB 3001 TTGCTCTCCGCTGGATATAAATGATATATTCAGGTAAAGGATATTCCTTATTTTCAGA 3060
QY 4367 CAAAAACATACACCACTGATGATCGAGGTGCAAGAGCATTTGCCCATGATCGAGAG 4426
DB 3061 CAAAAACATACACCACTGATGATCGAGGTGCAAGAGCATTTGCCCATGATCGAGAG 3120
QY 4427 GTCTTCCATCTGAACTGGAGGTCTTGTAAAAAATCTTTTCACTTCTGCTACGACAGAGA 4486
DB 3121 GTCTTCCATCTGAACTGGAGGTCTTGTAAAAAATCTTTTCACTTCTGCTACGACAGAGA 3180
QY 4487 TGAGTCTTACGSAATAATGAATTCATTCATTCAGATGCCCTCTCTGTTTTTAAAGAGTA 4546
DB 3181 TGAGTCTTACGSAATAATGAATTCATTCATTCAGATGCCCTCTCTGTTTTTAAAGAGTA 3240
QY 4547 CAGAGACACTGCACTACAGGAACATACAGATGATAACTGCCCTTTATTTGTCTGTATTG 4606
DB 3241 CAGAGACACTGCACTACAGGAACATACAGATGATAACTGCCCTTTATTTGTCTGTATTG 3300
QY 4607 AAATTTCTGGGTTTTCCAGCCAGCAACCAACTGAGTGCAATATGTAGTCATTCAGACTTTC 4666
DB 3301 AAATTTCTGGGTTTTCCAGCCAGCAACCAACTGAGTGCAATATGTAGTCATTCAGACTTTC 3360
QY 4667 AAGATATTCATATTCATGTTGGTGTGAGCAGACTATCCATAATCCTTTAAGAGTGTTC 4726
DB 3361 AAGATATTCATATTCATGTTGGTGTGAGCAGACTATCCATAATCCTTTAAGAGTGTTC 3420
QY 4727 CCTCTAAGTTTTCCGGGATTTCTGGATGCACTGATGGGGTCTCAAGAGGCTCAGCTA 4786
DB 3421 CCTCTAAGTTTTCCGGGATTTCTGGATGCACTGATGGGGTCTCAAGAGGCTCAGCTA 3480
QY 4787 GCAGCACCAAAAGCAGCAAGATTTGTTACTAGGTGTTAAAACAAATTCAGAGATGATACACAA 4846
DB 3481 GCAGCACCAAAAGCAGCAAGATTTGTTACTAGGTGTTAAAACAAATTCAGAGATGATACACAA 3540
QY 4847 TGTGCCGTATPACTCTTCCGCAAAAGAGTTCTAAAGATTAGTCATTAATTTGAGTAGTTTCAG 4906
DB 3541 TGTGCCGTATPACTCTTCCGCAAAAGAGTTCTAAAGATTAGTCATTAATTTGAGTAGTTTCAG 3600
QY 4907 TTTCAACTAAATGCTAGACTGGGCTTTTAAACAATTAAGGAGAGATTCCTCAAAACAT 4966
DB 3601 TTTCAACTAAATGCTAGACTGGGCTTTTAAACAATTAAGGAGAGATTCCTCAAAACAT 3660
QY 4967 TTTCACTGACATATGCTCTTACTCTGAGGTTTTCCCATTTGCTGTCACACTGCACATTCAGAC 5026
DB 3661 TTTCACTGACATATGCTCTTACTCTGAGGTTTTCCCATTTGCTGTCACACTGCACATTCAGAC 3720
QY 5027 TTTCCGCTGCGAGGTTTATCAAAAGAAATTTATTTCAAGATGTACAGTTTTCTACAAATGCATG 5086
DB 3721 TTTCCGCTGCGAGGTTTATCAAAAGAAATTTATTTCAAGATGTACAGTTTTCTACAAATGCATG 3780
QY 5087 AAGAGCAGAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTG 5146
DB 3781 AAGAGCAGAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTG 3840
QY 5147 AATCTCGACCTCATATTTATGATGGATATAGATACACTATATATATATATTTCTATTTCTGG 5206
DB 3841 AATCTCGACCTCATATTTATGATGGATATAGATACACTATATATATATTTCTATTTCTGG 3900
QY 5207 ATTTCTCTAAAAGCCTCAGAAAATACGACTGACTAGGAGCAAGAGAGAGTAGTATCTTCTG 5266
DB 3901 ATTTCTCTAAAAGCCTCAGAAAATACGACTGACTAGGAGCAAGAGAGAGTAGTATCTTCTG 3960
QY 5267 TACATGTTTCCGAGTTTACTGTTGATGATGAAACAGTTGGAACTGCTGACTTTCTTAACAAA 5326
DB 3961 TACATGTTTCCGAGTTTACTGTTGATGATGAAACAGTTGGAACTGCTGACTTTCTTAACAAA 4020
QY 5327 ACAACTTCTCTCTCTTCTTGTGAGCTTTTTCAGGGGTTTCATGATTCATTACACAGTT 5386
DB 4021 ACAACTTCTCTCTCTTCTTGTGAGCTTTTTCAGGGGTTTCATGATTCATTACACAGTT 4080
QY 5387 TTAAGAGTTTTCACTTACCAATTTGTATGCAAGAGCCAAAGCACTGAATACCTACATAGGTTTT 5446

Db 4081 TTAAGAGTTTCAGTTACCACTGATGACAGAGCAGACCTGATACCTACATAGTTT 4140
QY 5447 CTATTTTCTTCAATTTTAAAGCGTAATGACAGTGAACAATAATGGGATATGCAAGC 5506
Db 4141 CTATTTTCTTCAATTTTAAAGCGTAATGACAGTGAACAATAATGGGATATGCAAGC 4200
QY 5507 ACCCTTCACAGTTATTTCTGAATGA-TTTTAGGGTAAATAATACAGATGCCCTTGTATGT 5565
Db 4201 ACCCTTCACAGTTATTTCTGAATGAATTTTAGGGTAAATAATACAGATGCCCTTGTATGT 4260
QY 5566 TAACTAATCTGTGAAAGCAGGAATCAGTGTCTCTAAGGCTGCATCCTATTACCAAAAG 5625
Db 4261 TAACTAATCTGTGAAAGCAGGAATCAGTGTCTCTAAGGCTGCATCCTATTACCAAAAG 4320
QY 5626 GGGTTGTGCTATAACTGTGGCTGTATATAGAGAGGGAAC 5662
Db 4321 GGG-TGTGCTATAACT-GCTGGTATTAGAGAGGGAAC 4355

RESULT 4

ADQ89851
ID ADQ89851 standard; DNA; 3829 BP.
AC ADQ89851;
XX
XX
DT 21-OCT-2004 (first entry)
XX
DE Antagonist of cell cycle progression nucleotide sequence #141.
XX
XX Cytostatic; cancer; cell division cycle; mitosis; meiosis;
KW cell cycle progression; ds.
XX
XX Homo sapiens.
XX
XX WO2004063362-A2.
PN
PD 29-JUL-2004.
XX
XX 31-DEC-2003; 2003WO-GB005635.
PF
XX 10-JAN-2003; 2003US-0439123P.
PR 06-MAY-2003; 2003US-0468402P.
XX
XX (CYCL-) CYCLACEL LTD.
PA
XX
PI Glover D, Bell G, Frenz L, Midgley C;
XX
XX WPI; 2004-544089/52.
DR P-PSDB; ADQ89852.
XX

PT New cell cycle progression genes and proteins for modulating cell cycle
PT progression in cells, for preventing, treating or diagnosing cell
PT proliferative diseases (e.g. cancer) or for identifying modulators of
PT mitosis or meiosis.

Claim 1; SEQ ID NO 281; 461pp; English.

CC The present invention relates to a polynucleotide for preventing,
CC treating or diagnosing a disease in an individual. The composition or the
CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for
CC diagnosing, preventing or treating diseases (e.g. cell proliferative
CC diseases such as cancer) in an individual. These may also be used for
CC identifying substances capable of binding to or modulating the function
CC of the polypeptide, capable of affecting the function of the
CC corresponding gene, or capable of inhibiting the cell division cycle or
CC cell cycle progression, preferably mitosis and/or meiosis. The present
CC sequence represents an antagonist of cell cycle progression nucleotide
CC sequence.

Sequence 3829 BP; 1230 A; 776 C; 748 G; 1075 T; 0 U; 0 Other;

Query Match 67.6%; Score 3829; DB 13; Length 3829;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 3829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1326 GGCAACAACAATTTCTTCTCATTTACATAGCCATCATTTACACTGCTTGCCCAACCTTAAT 1385
Db 1 GGCAACAACAATTTCTTCTCATTTACATAGCCATCATTTACACTGCTTGCCCAACCTTAAT 60
QY 1386 GAATATGGCTGCATCCTTTGTATATCCCAAGAAAGAGACTGCGAGCCAGTGCAGCCTT 1445
Db 61 GAATATGGCTGCATCCTTTGTATATCCCAAGAAAGAGACTGCGAGCCAGTGCAGCCTT 120
QY 1446 GAACTGTTTAAAGCGTTCCTCATGAATGAAGAAAGAGACTGCGAGCCAGTGCAGCCTT 1505
Db 121 GAACTGTTTAAAGCGTTCCTCATGAATGAAGAAAGAGACTGCGAGCCAGTGCAGCCTT 180
QY 1506 TTTAGACCACATTAATTCAGAAAGCAATTCGAACACACACAGAAACCGGATCAGTATCTCG 1565
Db 181 TTTAGACCACATTAATTCAGAAAGCAATTCGAACACACACAGAAACCGGATCAGTATCTCG 240
QY 1566 AGTTCAGAAAGATATTTATCTTTAAGGATACAGAGGAAGCTCTTTTAAATTAACCTTAG 1625
Db 241 AGTTCAGAAAGATATTTATCTTTAAGGATACAGAGGAAGCTCTTTTAAATTAACCTTAG 300
QY 1626 AGATAGCCCAAGTCTTCAACATAAAGAGAACTTGAATGGAATTCGAATCTTATAGGAC 1685
Db 301 AGATAGCCCAAGTCTTCAACATAAAGAGAACTTGAATGGAATTCGAATCTTATAGGAC 360
QY 1686 CATCTTAAAGTGGCCAAATGTAAATCTAAGAACTATATAAGATGAACAGTTCACAGGTT 1745
Db 361 CATCTTAAAGTGGCCAAATGTAAATCTAAGAACTATATAAGATGAACAGTTCACAGGTT 420
QY 1746 TGTACGAAGACTACTTTATTTTACAGCCAGCAGTAAATTAATATGCAACCTGGAATCT 1805
Db 421 TGTACGAAGACTACTTTATTTTACAGCCAGCAGTAAATTAATATGCAACCTGGAATCT 480
QY 1806 GGATTTTGC CAAGGCCAAACAGCTCAGGTTGTAGTTGCCAGTTTACAGAAATTTCTTCT 1865
Db 481 GGATTTTGC CAAGGCCAAACAGCTCAGGTTGTAGTTGCCAGTTTACAGAAATTTCTTCT 540
QY 1866 TGAATCTGAAGAGGATGGCAAGGCTACTTGAAGATCTAGTAAAGATATTTGTTCACTG 1925
Db 541 TGAATCTGAAGAGGATGGCAAGGCTACTTGAAGATCTAGTAAAGATATTTGTTCACTG 600
QY 1926 GCTCAATGCTTCATCTGGAATGAACCGGAAAGAGCTTCAAAATTAATGTTTATTGAC 1985
Db 601 GCTCAATGCTTCATCTGGAATGAACCGGAAAGAGCTTCAAAATTAATGTTTATTGAC 660
QY 1986 CACCTTAGTCAACACTACTTTTATTTATTTGGAACACTTTCTTGCCACCCCTCATGGAGT 2045
Db 661 CACCTTAGTCAACACTACTTTTATTTATTTGGAACACTTTCTTGCCACCCCTCATGGAGT 720
QY 2046 TAAATGCTGAAATAATGCACTGATTTTCACTGCTCTTAAATCTTTGCTCTCTTGAATAA 2105
Db 721 TAAATGCTGAAATAATGCACTGATTTTCACTGCTCTTAAATCTTTGCTCTCTTGAATAA 780
QY 2106 CCAAGATCACTTGCTTAAACTTACTGTTTCTAGCTTGGACTATAGCAGAGATGATGGC 2165
Db 781 CCAAGATCACTTGCTTAAACTTACTGTTTCTAGCTTGGACTATAGCAGAGATGATGGC 840
QY 2166 TAGAGTCACTCTTTTCCAAAATTTTAACTGCAAGTACTGATGCTGCGAGACTTATGCAAC 2225
Db 841 TAGAGTCACTCTTTTCCAAAATTTTAACTGCAAGTACTGATGCTGCGAGACTTATGCAAC 900
QY 2226 AAAACATTTAAGGATATTTATGAGAGCTAATGTTGAATCTCTTTAATTAATGGGAAATGA 2285
Db 901 AAAACATTTAAGGATATTTATGAGAGCTAATGTTGAATCTCTTTAATTAATGGGAAATGA 960
QY 2286 GTTCTTAGTGAACCCAGCTACATGATATAAAACAAACGATTTCTCTGAGCTCTTGATAT 2345
Db 961 GTTCTTAGTGAACCCAGCTACATGATATAAAACAAACGATTTCTCTGAGCTCTTGATAT 1020
QY 2346 CCTCGATGAAGCATGTGAAGCAAGGCAATCTTCATGCTCTCATTCAGATGAACCAAGC 2405

Db 1021 CCTGATGAAGCATGTGAAGACAGGCCAACTTTCATGCTCTCATTTAGATGAACCCAGC 1080
Qy 2406 GTTATCCCACTTGGAGACAAGGGTTTGCCTTCTCTGCTGAGATTTCTCTCCATTCCAAA 2465
Db 1081 GTTATCCCACTTGGAGACAAGGGTTTGCCTTCTCTGCTGAGATTTCTCTCCATTCCAAA 1140
Qy 2466 AGGATTTTCCCTATCTGAATGAAGAGGTATGTAGCAAAACAATTTGGAAAAAGTGGCAG 2525
Db 1141 AGGATTTTCCCTATCTGAATGAAGAGGTATGTAGCAAAACAATTTGGAAAAAGTGGCAG 1200
Qy 2526 GGAATACAACTCCAAATATGTTGACTTCATTTGAGGAACAACCTCAATGAAGCACTTTACTAC 2585
Db 1201 GGAATACAACTCCAAATATGTTGACTTCATTTGAGGAACAACCTCAATGAAGCACTTTACTAC 1260
Qy 2586 TTACCGGAAGCCTGTGTGATGATTAATTTACAGAACTCTGTCGTAATGTTTCGTACACAGA 2645
Db 1261 TTACCGGAAGCCTGTGTGATGATTAATTTACAGAACTCTGTCGTAATGTTTCGTACACAGA 1320
Qy 2646 TCCTCACGCTCTACCTGCTATACACCTTTATGGACAACCTAGTACACCAATAAAACAGGCTG 2705
Db 1321 TCCTCACGCTCTACCTGCTATACACCTTTATGGACAACCTAGTACACCAATAAAACAGGCTG 1380
Qy 2706 CCATTTGTTGGAAGTACAGAAATATTTATACAGAACTCTGTCGTAATGTTTCGTACACAGA 2765
Db 1381 CCATTTGTTGGAAGTACAGAAATATTTATACAGAACTCTGTCGTAATGTTTCGTACACAGA 1440
Qy 2766 TTTGGATAAGTGGGAAGAAATTTAAAAAACTGAAAGCATCTCTTTGGGCTTTGGGAATAT 2825
Db 1441 TTTGGATAAGTGGGAAGAAATTTAAAAAACTGAAAGCATCTCTTTGGGCTTTGGGAATAT 1500
Qy 2826 CGGCTCATCAAAATTTGGGCTCAATTTGCTACAGGAAGAAACGTGATTCAGATATACT 2885
Db 1501 CGGCTCATCAAAATTTGGGCTCAATTTGCTACAGGAAGAAACGTGATTCAGATATACT 1560
Qy 2886 AAAAATTGCAAAAAGTGTGAAGTCTTTCCATCAGAGGACCTGTGTATGATGACTTGG 2945
Db 1561 AAAAATTGCAAAAAGTGTGAAGTCTTTCCATCAGAGGACCTGTGTATGATGACTTGG 1620
Qy 2946 GCTCATAGTAAAAACCAAGGCTGTGATTTCTAAAAATGTCACAACTGGGATGCTGT 3005
Db 1621 GCTCATAGTAAAAACCAAGGCTGTGATTTCTAAAAATGTCACAACTGGGATGCTGT 1680
Qy 3006 GAGGATAGTCGCAAAACATCTGTGGCCAGTGGTTCAGATGATGTGGAACAACCTCTGTAA 3065
Db 1681 GAGGATAGTCGCAAAACATCTGTGGCCAGTGGTTCAGATGATGTGGAACAACCTCTGTAA 1740
Qy 3066 TGAACCTTTCATCTATCCCAAGCACTCTAAGTTTGAATCTCGAGTCAACCACTCTAGACA 3125
Db 1741 TGAACCTTTCATCTATCCCAAGCACTCTAAGTTTGAATCTCGAGTCAACCACTCTAGACA 1800
Qy 3126 TAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTCAATTTGGAGGATGACCGTTTGG 3185
Db 1801 TAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTCAATTTGGAGGATGACCGTTTGG 1860
Qy 3186 CAGCAGCTCTACTAGTACATTTTCTTTGATPATCAATGAAGATACAGAGCCAAACATTTTA 3245
Db 1861 CAGCAGCTCTACTAGTACATTTTCTTTGATPATCAATGAAGATACAGAGCCAAACATTTTA 1920
Qy 3246 TGACCGATCTGGACCCATAAAGGATAAAAATTCATTCCTCTTCTTCTTCTAGTAACT 3305
Db 1921 TGACCGATCTGGACCCATAAAGGATAAAAATTCATTCCTCTTCTTCTTCTAGTAACT 1980
Qy 3306 TGTGAAGATCGTATCTTAAATTCGCTTACCTTTGCTTAAACAAAAACATCGTACTAGCAG 3365
Db 1981 TGTGAAGATCGTATCTTAAATTCGCTTACCTTTGCTTAAACAAAAACATCGTACTAGCAG 2040
Qy 3366 TGATCCAAAAGGAGGGAATTTATCATCTGAAGTAAGAACAAGCAACGAGGGAATCAGAAC 3425
Db 2041 TGATCCAAAAGGAGGGAATTTATCATCTGAAGTAAGAACAAGCAACGAGGGAATCAGAAC 2100
Qy 3426 ACTTACGAGCCCAAGTGTGATTTTAAATCATAGTGAATTTTACCCATATCCACTGT 3485
Db 2101 ACTTACGAGCCCAAGTGTGATTTTAAATCATAGTGAATTTTACCCATATCCACTGT 2160

Qy 3486 ACAGAAAACATTACAATTTAGAGACTTTCATTTATGGGAATAAGCACATTTGAAGACACTGG 3545
Db 2161 ACAGAAAACATTACAATTTAGAGACTTTCATTTATGGGAATAAGCACATTTGAAGACACTGG 2220
Qy 3546 TAGTACACCAAGCAATGGAGAAAATGACTTAAATTTCAACCAAGAAATTTTGGTACAGAGAA 3605
Db 2221 TAGTACACCAAGCAATGGAGAAAATGACTTAAATTTCAACCAAGAAATTTTGGTACAGAGAA 2280
Qy 3606 TCACAGAGAAAATCAACAGCCGAGAGAGTTAGTAGTAGAAAAGTTCAACGAGCTCACATAT 3665
Db 2281 TCACAGAGAAAATCAACAGCCGAGAGAGTTAGTAGTAGAAAAGTTCAACGAGCTCACATAT 2340
Qy 3666 GAAGATACGTAGCCAAAGTTTCAATACAGACACTACAACAGTGGCATAAGTTCAATGAG 3725
Db 2341 GAAGATACGTAGCCAAAGTTTCAATACAGACACTACAACAGTGGCATAAGTTCAATGAG 2400
Qy 3726 CTCAAGTCTTTCACAGAGACAGTAGGTGTAGATGCTACAACATATGGACACAGACTGTGG 3785
Db 2401 CTCAAGTCTTTCACAGAGACAGTAGGTGTAGATGCTACAACATATGGACACAGACTGTGG 2460
Qy 3786 AAGCATGAGTACTGTGTAAAGTACTTAAACCTATTAAAGCAAGCCACTATTTGACGCCACA 3845
Db 2461 AAGCATGAGTACTGTGTAAAGTACTTAAACCTATTAAAGCAAGCCACTATTTGACGCCACA 2520
Qy 3846 GTCTAAACCATCTCTCTCTCCAAATCAAAATCCGTTGTCCTGCTGCCAGGTTCTTC 3905
Db 2521 GTCTAAACCATCTCTCTCTCCAAATCAAAATCCGTTGTCCTGCTGCCAGGTTCTTC 2580
Qy 3906 TCATACGCTTCTTAGAAGAGACAGTCCCTTAAAGCACCTCTATTTGCTACAATTAAGAAG 3965
Db 2581 TCATACGCTTCTTAGAAGAGACAGTCCCTTAAAGCACCTCTATTTGCTACAATTAAGAAG 2640
Qy 3966 TCTAGCAGATGTAACTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGCTATGCTACACT 4025
Db 2641 TCTAGCAGATGTAACTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGCTATGCTACACT 2700
Qy 4026 GAAAGACTACAGCAACAAAGAAATGCATCCATCCCTATCTCACCTGAAGCTTTGGGATC 4085
Db 2701 GAAAGACTACAGCAACAAAGAAATGCATCCATCCCTATCTCACCTGAAGCTTTGGGATC 2760
Qy 4086 TCCAGCAAAAGATGTGCTATTTACTGATACCATCAATGAAGCCCAACAGTTTGTAGTC 4145
Db 2761 TCCAGCAAAAGATGTGCTATTTACTGATACCATCAATGAAGCCCAACAGTTTGTAGTC 2820
Qy 4146 CAGATTAACACCAAGCAGGTTCAAGAGCTTAAAGTTATGCAATCATTTAGATTAAGAGAA 4205
Db 2821 CAGATTAACACCAAGCAGGTTCAAGAGCTTAAAGTTATGCAATCATTTAGATTAAGAGAA 2880
Qy 4206 TTTATTGAGTCTTATTAAATCAAAATACCTGCAACGATCTTCCCTCAGTGGCGTCCATGGT 4265
Db 2881 TTTATTGAGTCTTATTAAATCAAAATACCTGCAACGATCTTCCCTCAGTGGCGTCCATGGT 2940
Qy 4266 GTCCAGTGCCACATATGGGGTTTCAGATGATTAATTTGCTCTTCCCGGTGGATAT 4325
Db 2941 GTCCAGTGCCACATATGGGGTTTCAGATGATTAATTTGCTCTTCCCGGTGGATAT 3000
Qy 4326 AAATGATATATTTCCAGTAAAGGATATTCCTATTTTTCAGACAAAAACAATACCAACACA 4385
Db 3001 AAATGATATATTTCCAGTAAAGGATATTCCTATTTTTCAGACAAAAACAATACCAACACA 3060
Qy 4386 TGATGATCGAGGTGCAAGAGCATTTTGGCCATGATGCAAGGAGTCTTCCATCTGGAACCTGG 4445
Db 3061 TGATGATCGAGGTGCAAGAGCATTTTGGCCATGATGCAAGGAGTCTTCCATCTGGAACCTGG 3120
Qy 4446 AGGTCTTTGAAAAAATTTCTTTTCACTTGTCTACGACAGCAGATGAGTCTTACGGAAATAAT 4505
Db 3121 AGGTCTTTGAAAAAATTTCTTTTCACTTGTCTACGACAGCAGATGAGTCTTACGGAAATAAT 3180
Qy 4506 GAATTCATTCATTCAGATGCTCTCTGTTTATAGAAAGTACAGAAAGACACTGAGCTACA 4565
Db 3181 GAATTCATTCATTCAGATGCTCTCTGTTTATAGAAAGTACAGAAAGACACTGAGCTACA 3240

QY 4566 GGAACATACAGATGATAACGCTTTATTTGCTGTATGAAATCTTGGGTTTCCAGCC 4625
Db |||||
QY 3241 GGAACATACAGATGATAACGCTTTATTTGCTGTATGAAATCTTGGGTTTCCAGCC 3300
Db |||||
QY 4626 CAGCAACCACTGAGTGCATATGTAGTCAATTCAGACTTTCAAGATATTCATATTCGA 4685
Db |||||
QY 3301 CAGCAACCACTGAGTGCATATGTAGTCAATTCAGACTTTCAAGATATTCATATTCGA 3360
Db |||||
QY 4686 TTGGTGTGACGACACTATCCATATCTTTAGAGTGGTTCCTCTAAAGTTTTCGGGGAT 4745
Db |||||
QY 3361 TTGGTGTGACGACACTATCCATATCTTTAGAGTGGTTCCTCTAAAGTTTTCGGGGAT 3420
Db |||||
QY 4746 TTCTGGATGACGATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCACCAAAAGACAGA 4805
Db |||||
QY 3421 TTCTGGATGACGATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCACCAAAAGACAGA 3480
Db |||||
QY 4806 ATTGTTACTAGGTGTTAAACAATTCAGATGATACACCAATGTGCGGTATCTCTTCG 4865
Db |||||
QY 3481 ATTGTTACTAGGTGTTAAACAATTCAGATGATACACCAATGTGCGGTATCTCTTCG 3540
Db |||||
QY 4866 CAAAGAAGTTCTAAGATTAGTCAATTAATTTGAGTGTTCAGTTTCAACTAAATGTCTGA 4925
Db |||||
QY 3541 CAAAGAAGTTCTAAGATTAGTCAATTAATTTGAGTGTTCAGTTTCAACTAAATGTCTGA 3600
Db |||||
QY 4926 GACTGGGCTTTTAAACAATTAAGAGAGATATCTCTCAACAATTTGATGACATATGCTTTA 4985
Db |||||
QY 3601 GACTGGGCTTTTAAACAATTAAGAGAGATATCTCTCAACAATTTGATGACATATGCTTTA 3660
Db |||||
QY 4986 CTCTGAGGTTTCCATTTGCTGTCACATGTCACATTCAGACTTCCGTGCGAGGTTTCAT 5045
Db |||||
QY 3661 CTCTGAGGTTTCCATTTGCTGTCACATGTCACATTCAGACTTCCGTGCGAGGTTTCAT 3720
Db |||||
QY 5046 ACAAGATATTTTCAAGATGTACAGTTTCTTACAAATGCAATGCAAGACGAGGCTGTGT 5105
Db |||||
QY 3721 ACAAGATATTTTCAAGATGTACAGTTTCTTACAAATGCAATGCAAGACGAGGCTGTGT 3780
Db |||||
QY 5106 GGCAACACCAACCAACCACTATAGTTGATATCTGCTGAATCCTGA 5154
Db |||||
QY 3781 GGCAACACCAACCAACCACTATAGTTGATATCTGCTGAATCCTGA 3829
Db |||||

RESULT 5

ADC30081

ID ADC30081 standard; cDNA; 2659 BP.

XX AC ADC30081;

XX AC ADC30081;

XX DT 18-DEC-2003 (first entry)

XX DE Human novel cDNA sequence, SEQ ID NO:163.

XX KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 5; gene; ss.

XX OS Homo sapiens.

XX PN WO2003029271-A2.

XX XX 10-APR-2003.

XX PD 24-SEP-2002; 2002WO-US030474.

XX PF 24-SEP-2001; 2001US-0324631P.

XX PR (HYSE-) HYSEQ INC.

XX PA Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX XX

PI

PI

XX

DR

XX

PT

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

Halley-Vicente D, Drmanac RT;

WPI; 2003-371981/35.

P-PSDB; ADC31052.

New polynucleotide and polypeptide useful for diagnosing, preventing or

treating conditions such as neurodegenerative diseases, anemias, platelet

disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

cancer.

Claim 1; SEQ ID NO 163; 1185pp; English.

The invention relates to 971 novel human cDNA sequences (ADC39919-

ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

invention also relates to nucleic acid sequences over 99% identical with

the novel human cDNAs. The invention additionally encompasses expression

vectors and host cells comprising a nucleic acid of the invention; the

recombinant production of a polypeptide of the invention; an antibody

against a polypeptide of the invention; a method of detecting of

polynucleotides or polypeptides of the invention; and methods of

identifying a compound which binds to a polypeptide of the invention. The

invention further discloses methods of patenting, treating or

ameliorating a medical condition; kits comprising polynucleotide probes

and/or monoclonal antibodies for carrying out the methods of the

invention; methods for the identification of compounds that modulate the

expression or activity of the polynucleotide and/or polypeptide; and 767

config sequences corresponding to the cDNA sequences of the invention

(ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628

-ADC33394). The nucleic acids and polypeptides of the invention are

useful in diagnostics, drug screening, forensics, gene mapping, in the

identification of mutations responsible for genetic disorders or other

traits, for assessing biodiversity, and in producing many other types of

data and products dependent on DNA and amino acid sequences. They are

also used for treating diseases such as Parkinson's disease, Alzheimer's

disease and other neurodegenerative diseases, anaemia, platelet

disorders, wounds, ulcers, osteoporosis, autoimmune diseases or

cancer. The nucleic acids may also be used as hybridisation probes or

primers, and in the recombinant production of a protein. The polypeptides

are also useful in generating antibodies, as molecular weight markers,

and as food supplements. The present sequence represents a specifically

claimed human cDNA sequence of the invention. Note: The sequence data for

this patent did not form part of the printed specification, but was

obtained in electronic format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences.

Sequence 2659 BP; 848 A; 539 C; 524 G; 748 T; 0 U; 0 Other;

Query Match 40.2%; Score 2276.4; DB 10; Length 2659;

Best Local Similarity 95.9%; Pred. No. 0;

Matches 2386; Conservative 0; Mismatches 6; Indels 96; Gaps 1;

QY 2917 ATCAGAGGACCTGTGTATATGTACTTGGGCTCATAGCTAAACCAACAGGCTGTGAT 2976

Db 268 AACCAAGAGACCTGTGTATATGTACTTGGGCTCATAGCTAAACCAACAGGCTGTGAT 327

QY 2977 ATTCTAAATGTCAACACTGGGATCTGTGAGGCTAGTCGCAACATCTGGCCAGTG 3036

Db 328 ATTTAAATGTCAACACTGGGATCTGTGAGGCTAGTCGCAACATCTGGCCAGTG 387

QY 3037 GTTCCAGATGTGTGGAACAACTCTGTATGAATCTATCTATCCAGCACTCTAAGT 3096

Db 388 GTTCCAGATGTGTGGAACAACTCTGTATGAATCTATCTATCCAGCACTCTAAGT 447

QY 3097 TTGAACTCGAGTCAACAGGCTCTAGACATATAGTGAAGTGAATCTGTGCCATCGAGT 3156

Db 448 TTGAACTCGAGTCAACAGGCTCTAGACATATAGTGAAGTGAATCTGTGCCATCGAGT 507

QY 3157 ATGTTTCATATTTGGAGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTTTCTCTGAT 3216

Db 508 ATGTTTCATATTTGGAGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTTTCTCTGAT 567

QY 3217 ATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCCAATGAAGATAAAAT 3276

[illegible]

RESULT 6
ADL45178/C
ID ADL45178 standard; DNA; 4657 BP.
XX
AC ADL45178;
XX
XX 20-MAY-2004 (first entry)
DT
XX
DE Human ovarian cancer DNA marker #19068.
XX
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
XX Homo sapiens.
OS
XX WO200170979-A2.
PN
XX
PD 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009126.
PF
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX Lee J, Lillie J;
PI
XX WPI; 2001-611502/70.
DR
XX
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
PT
XX
XX Disclosure; SEQ ID NO 19068; 106pp; English.
PS
XX
XX The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX

SQ	Sequence	4657 BP; 1440 A; 860 C; 877 G; 1476 T; 0 U; 4 Other;
	Query Match	39.6%; Score 2244.2; DB 5; Length 4657;
	Best Local Similarity	95.9%; Pred. No. 0;
	Matches 2386; Conservative	0; Mismatches 3; Indels 100; Gaps 4;
QY	3175	GAACGGTTTGGCAGCAGCTCTACTAGTACATTTTCTTGTATATCAATGAAGATACAGAG 3234
DB	4634	GACCGGTTTGGCAGCAGCTCTACTAGTACATTTTCTTGTATATCAATGAAGATACAGAG 4575
QY	3235	CCAAATTTTATGACCGATCTGCACCCATAAAGGATAAAATTCATTCCTCTTCTTGTCT 3294
DB	4574	CCAAATTTTATGACCGATCTGCACCCATAAAGGATAAAATTCATTCCTCTTCTTGTCT 4515
QY	3295	TCTAGTAAACTTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCCTAACAAAAACAT 3354
DB	4514	TCTAGTAAACTTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCCTAACAAAAACAT 4455
QY	3355	CGTAGTAGCAGTGATCCCAAAAGGAGGAATATCATCTGAAAGTAAGCAAGCAACAGG 3414
DB	4454	CGTAGTAGCAGTGATCCCAAAAGGAGGAATATCATCTGAAAGTAAGCAAGCAACAGG 4395
QY	3415	CGAATCAGAAACACTTTACGGAGCCCACTGTTGATTTTAAATCATAGTGATGATTTTACACC 3474
DB	4394	CGAATCAGAAACACTTTACGGAGCCCACTGTTGATTTTAAATCATAGTGATGATTTTACACC 4335
QY	3475	ATATCCACTGTACAGAAAACATTTACAATTAGAGACTTCATTTATGCGGAATAAGCACATT 3534
DB	4334	ATATCCACTGTACAGAAAACATTTACAATTAGAGACTTCATTTATGCGGAATAAGCACATT 4275
QY	3535	GAAGACACTGGTAGTACACCAAGCAATTCGAGAGAAATGACTTAAATTTCCCAAGAAATTTT 3594
DB	4274	GAAGACACTGGTAGTACACCAAGCAATTCGAGAGAAATGACTTAAATTTCCCAAGAAATTTT 4215
QY	3595	GGTACAGAGAATCAACAGAGAAAATAACAAGCCGAGAGAGTTAGTAGTAAAGTTTCAACG 3654
DB	4214	GGTACAGAGAATCAACAGAGAAAATAACAAGCCGAGAGAGTTAGTAGTAAAGTTTCAACG 4155
QY	3655	AGCTCACATNTGAAGATAGGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 3714
DB	4154	AGCTCACATNTGAAGATAGGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 4095
QY	3715	AGTTCAATGAGCTCAAGTCCTTTCAGAGAGACAGTAGGTGTAGATGCTTACAACTATGAC 3774
DB	4094	AGTTCAATGAGCTCAAGTCCTTTCAGAGAGACAGTAGGTGTAGATGCTTACAACTATGAC 4035
QY	3775	ACAGACTGTGAAGCATGAGTACTGTGTAAAGTAACTACTAAAGTAAAGCAAGCCACTAT 3834
DB	4034	ACAGACTGTGAAGCATGAGTACTGTGTAAAGTAACTACTAAAGTAAAGCAAGCCACTAT 4009
QY	3835	TTGAGCCACAGCTCTAACCATCTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTGCT 3894
DB	4008	TTGAGCCACAGCTCTAACCATCTGTCTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTGCT 4009
QY	3895	CCAGGTTCTTCTCATAGCTTCTTAGAAGAGACAGTGCCTTAAAGCAACCTCTATTGCT 3954
DB	4008	CCAGGTTCTTCTCATAGCTTCTTAGAAGAGACAGTGCCTTAAAGCAACCTCTATTGCT 3952
QY	3955	ACAATTTAAAGTCTAGCAGATTTGAACTTTAGTTTACACAAAGTTCTAGAGATGCTTTGGC 4014
DB	3951	ACAATTTAAAGTCTAGCAGATTTGAACTTTAGTTTACACAAAGTTCTAGAGATGCTTTGGC 3892
QY	4015	TATGCTACACTGAAAGACTACACCAAGAAATGATCCATCTTATCTCTACTCTCTGAA 4074
DB	3891	TATGCTACACTGAAAGACTACACCAAGAAATGATCCATCTTATCTCTACTCTCTGAA 3832
QY	4075	GCTTTGGCATCTCCAGCAAAAGATGTGCTATTCTTACTGATACCATCAGCAAGCCCAAC 4134
DB	3831	GCTTTGGCATCTCCAGCAAAAGATGTGCTATTCTTACTGATACCATCAGCAAGCCCAAC 3772
QY	4135	AGTTTTCAGTCCAGATTTAACACCAAGAGGTTTCAATGAAGCCTTAAGTTATGATCATTTA 4194
DB	3771	AGTTTTCAGTCCAGATTTAACACCAAGAGGTTTCAATGAAGCCTTTAAGTTATGATCATTTA 3712

QY 4195 GATTAAGAGAGATTTATTGAGTCTTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTG 4254
DB 3711 GATAAAGAGAGATTTATTGAGTCTTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTG 3652
QY 4255 CGGTCCATGGTCCAGTGCACATATATGGGGTTTCAGATGATTAATGTTCTGCTCTC 4314
DB 3651 CGGTCCATGGTCCAGTGCACATATATGGGGTTTCAGATGATTAATGTTCTGCTCTC 3592
QY 4315 CCGTGGATATAATGATATATCCAGTAAAGGATATTCCTATTTTCAGACAAAAC 4374
DB 3591 CCGTGGATATAATGATATATCCAGTAAAGGATATTCCTATTTTCAGACAAAAC 3532
QY 4375 ATACCAACATGATGATCGAGTGCAGAGCAATTTGCCATGATGAGGAGTCTTCCA 4434
DB 3531 ATACCAACATGATGATCGAGTGCAGAGCAATTTGCCATGATGAGGAGTCTTCCA 3472
QY 4435 TCTGGAATCGAGTCTTGTAAAAATCTTTTCACTGCTACGACAGCAGATGAGTCTT 4494
DB 3471 TCTGGAATCGAGTCTTGTAAAAATCTTTTCACTGCTACGACAGCAGATGAGTCTT 3412
QY 4495 ACGGAATATGAAATCAATCCATTCAGATGCCCTCTCTGTTTTCAGAAAGTACAGAAGAC 4554
DB 3411 ACGGAATATGAAATCAATCCATTCAGATGCCCTCTCTGTTTTCAGAAAGTACAGAAGAC 3352
QY 4555 ACTGGAATACAGGAACATACAGATGATGATGCTCTTATTGTTGCTGTTTGAATTTCTG 4614
DB 3351 ACTGGAATACAGGAACATACAGATGATGATGCTCTTATTGTTGCTGTTTGAATTTCTG 3292
QY 4615 GGTTCAGCCAGCCAGCAACCACTGAGTGCAATATGTTAGTCAATTCAGACTTTCAAGATATT 4674
DB 3291 GGTTCAGCCAGCCAGCAACCACTGAGTGCAATATGTTAGTCAATTCAGACTTTCAAGATATT 3232
QY 4675 CCATATTCTGATGCTGAGCAGACTATCCATATCTTTAGAGTGGTTTCCCTCTAAG 4734
DB 3231 CCATATTCTGATGCTGAGCAGACTATCCATATCTTTAGAGTGGTTTCCCTCTAAG 3172
QY 4735 TTTTCGGGGATTTCTGGATGAGTATGAGTGGTGTCTCAAGAAGCTCAGCTAGCAGCAC 4794
DB 3171 TTTTCGGGGATTTCTGGATGAGTATGAGTGGTGTCTCAAGAAGCTCAGCTAGCAGCAC 3112
QY 4795 AAAAGCAGAGATTTGTTACTAGTGTAAACAAATTCAGATGATACCAATGTCGCGT 4854
DB 3111 AAAAGCAGAGATTTGTTACTAGTGTAAACAAATTCAGATGATACCAATGTCGCGT 3052
QY 4855 ATACTCTTCGCAAGAGTTCTAAGATTAGTCAATTAATTTGAGTGTTCAGTTTCAACT 4914
DB 3051 ATACTCTTCGCAAGAGTTCTAAGATTAGTCAATTAATTTGAGTGTTCAGTTTCAACT 2992
QY 4915 AAATGTCATGAGACTGGGCTTTTAAACAAATTAAGGAGAGTATCTCAAAATTTGATGAC 4974
DB 2991 AAATGTCATGAGACTGGGCTTTTAAACAAATTAAGGAGAGTATCTCAAAATTTGATGAC 2932
QY 4975 ATATGCTTTTACTGAGGTTTCCATTTGCTGTGCATGCAATTCAGACTTCCGTTGT 5034
DB 2931 ATATGCTTTTACTGAGGTTTCCATTTGCTGTGCATGCAATTCAGACTTCCGTTGT 2872
QY 5035 CGGAGTTTCATACAGAAATTTTTCAGATGATACAGTTTCTCAAAATGCAAGAGCA 5094
DB 2871 CGGAGTTTCATACAGAAATTTTTCAGATGATACAGTTTCTCAAAATGCAAGAGCA 2812
QY 5095 GAGGCTGTGTTGGCAACACCAACCAACCACTATAGTTGATACATCTGCTGAATCTGA 5154
DB 2811 GAGGCTGTGTTGGCAACACCAACCAACCACTATAGTTGATACATCTGCTGAATCTGA 2752
QY 5155 CCTCATATTTATGATGATATAGATACATATATATATATATATTTGTTGGAATTCCTA 5214
DB 2751 CCTCATATTTATGATGATATAGATACATATATATATATATATTTGTTGGAATTCCTA 2692
QY 5215 AAAGCTCAGAAATACGACTGAGGAGCAAGCAGGATATCTTCTGACTGT 5274
DB 2691 AAAGCTCAGAAATACGACTGAGGAGCAAGCAGGATATCTTCTGACTGT 2632

QY 5275 TCCGCACTTACTGGTATCATGAACAGATTGGAACTGCTGACTTTCTTAACCAACAACTTC 5334
DB 2631 TCCGCACTTACTGGTATCATGAACAGATTGGAACTGCTGACTTTCTTAACCAACAACTTC 2572
QY 5335 CTCTCTCCCTTTGTTGAGCCCTTTTGGGGTTTCATGATTCATTACCAAGTTTAAAGT 5394
DB 2571 CTCTCTCCCTTTGTTGAGCCCTTTTGGGGTTTCATGATTCATTACCAAGTTTAAAGT 2512
QY 5395 TTCACTTTTAAACCGTAAATGACAGTGGAACTAATTAATGGATATGCAAGACACCTTCA 5454
DB 2511 TTCACTTTTAAACCGTAAATGACAGTGGAACTAATTAATGGATATGCAAGACACCTTCA 2452
QY 5455 TTTCACTTTTAAACCGTAAATGACAGTGGAACTAATTAATGGATATGCAAGACACCTTCA 5514
DB 2451 TTTCACTTTTAAACCGTAAATGACAGTGGAACTAATTAATGGATATGCAAGACACCTTCA 2392
QY 5515 CAAGTTATTTCTGAATGA-TTTTAGGGTAAATATACAGATGCCCTTGTATGTTAACTAAC 5573
DB 2391 CAAGTTATTTCTGAATGA-TTTTAGGGTAAATATACAGATGCCCTTGTATGTTAACTAAC 2332
QY 5574 TTGTGGAAGCAGCAATCAGTGTCTTAAGGCTGCATCTTATACCAATGGGTTGTG 5633
DB 2331 TTGTGGAAGCAGCAATCAGTGTCTTAAGGCTGCATCTTATACCAATGGG-TGTG 2273
QY 5634 CTATACTGGCTGCTATTAGAGAGGGAAC 5662
DB 2272 CTATACT-GCTGCTATTAGAGAGGGAAC 2245

RESULT 7
ABV20190/c
ID ABV20190 standard; cDNA; 4601 BP.
XX
AC ABV20190;
XX
XX 13-SEP-2002 (first entry)
DT
XX Human prostate expression marker cDNA 20181.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; Gene; SS.
XX Homo sapiens.
XX
XX WO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3296; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 4601 BP; 1416 A; 852 C; 871 G; 1460 T; 0 U; 2 Other;

Query Match 39.6%; Score 2239.4; DB 5; Length 4601;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 2383; Conservative 0; Mismatches 6; Indels 100; Gaps 4;

QY 3175 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCCCTTGATATCAATGAAGTACAGAG 3234
DB 4580 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCCCTTGATATCAATGAAGTACAGAG 4521
QY 3235 CCAACATTTTATGACCGATCTGACCCATAAAGGATAAAAAATTCATTCCTTTCTTTGCT 3294
DB 4520 CCAACATTTTATGACCGATCTGACCCATAAAGGATAAAAAATTCATTCCTTTCTTTGCT 4461
QY 3295 TCTAGTAACTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCCTAAACAAAAACAT 3354
DB 4460 TCTAGTAACTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCCTAAACAAAAACAT 4401
QY 3355 CGTGTAGCAGTGTCCAAAGGAGGAAATTCATCTGAAAGTAAGACAGCAGCAGG 3414
DB 4400 CGTGTAGCAGTGTCCAAAGGAGGAAATTCATCTGAAAGTAAGACAGCAGCAGG 4341
QY 3415 CGAATCAGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACC 3474
DB 4340 CGAATCAGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACC 4281
QY 3475 ATATCCACTGTACAGAAAACTTACAAATPAGAGCTTCAATTTATGGGGAATPAGCACA 3534
DB 4280 ATATCCACTGTACAGAAAACTTACAAATPAGAGCTTCAATTTATGGGGAATPAGCACA 4221
QY 3535 GAAGACACTGGTAGTACACAGCAATGGAGAAATGACTTAAATTCACCAAGAAATTT 3594
DB 4220 GAAGACACTGGTAGTACACCAAGCAATGGAGAAATGACTTAAATTCACCAAGAAATTT 4161
QY 3595 GGTACAGAGAATCACAGAGAAAAATCAAGCCAGAGAGGTTAGTAGTAAAGTTCAACG 3654
DB 4160 GGTACAGAGAATCACAGAGAAAAATCAAGCCAGAGAGGTTAGTAGTAAAGTTCAACG 4101
QY 3655 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 3714
DB 4100 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 4041
QY 3715 AGTTCAATGAGCTCAAGCTCTTACAGAGAGACAGTAGGTGATGCTACAACTATGGAC 3774
DB 4040 AGTTCAATGAGCTCAAGCTCTTACAGAGAGACAGTAGGTGATGCTACAACTATGGAC 3981
QY 3775 ACAGACTGTGGAAGCATGAGTACTGTGGTAAAGTAACTTAAAGACAGCAGCACTAT 3834
DB 3980 ACAGACTGTGGAAGCATGAGTACTGTGGTAAAGTAACTTAAAGACAGCAGCACTAT 3955
QY 3835 TTGACGCCACAGCTAAACCATCTGTCTCTCTCCAAATCAAATTCGGTGTGCGCT 3894
DB 3954 ----- 3955
QY 3895 CCAGGTTCTTCTCATACGCTTCTAGAGAGACAGTCCCTTAAAGCACCCTCTATTGCT 3954
DB 3954 ---GGTTCTTCGTCAGGCTTCTTGAAGAGACAGTCCCTTAAAGCACCCTCTATTGCT 3898
QY 3955 ACAATTTAAAGCTAGCAGATTGTAATCTTAGTTACACAAAGTTCTAGAGATGCTTTTGC 4014
DB 3897 ACAATTTAAAGCTAGCAGATTGTAATCTTAGTTACACAAAGTTCTAGAGATGCTTTTGC 3838
QY 4015 TATGTCTACTGAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTCTGAA 4074
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 3837 TATGTCTACTGAAAGACTACAGCAACAAAGATGTCATCTCTTCTCTCACTCTGAA 3778
QY 4075 GCTTTGGCATCTCCAGCAAAAGATGTGTATTTACTGATACCATCACCATGAGGCCAAC 4134
DB 3777 GCTTTGGCATCTCCAGCAAAAGATGTGTATTTACTGATACCATCACCATGAGGCCAAC 3718
QY 4135 AGTTTGTAGTCCAGATTAACACCAAGCAGGTTTCATGAAGCCTTAAGTTATGCATCATTA 4194
DB 3717 AGTTTGTAGTCCAGATTAACACCAAGCAGGTTTCATGAAGCCTTAAGTTATGCATCATTA 3658
QY 4195 GATAAAGAGATTTTATTCAGTCTTATTAATCAAAATACCTGCAACGATCTTCTCAGTG 4254
DB 3657 GATAAAGAGATTTTATTCAGTCTTATTAATCAAAATACCTGCAACGATCTTCTCAGTG 3598
QY 4255 CGGTCCATGGTGTCCAGTCCACATATGGGGTTCAGATGATTACATTTGGTCTTCTCTC 4314
DB 3597 CGGTCCATGGTGTCCAGTCCACATATGGGGTTCAGATGATTACATTTGGTCTTCTCTC 3538
QY 4315 CCGGTGGATATAAATGATATATTCAGGTAAAGGATATTCCTATTTTTCAGACAAAAAC 4374
DB 3537 CCGGTGGATATAAATGATATATTCAGGTAAAGGATATTCCTATTTTTCAGACAAAAAC 3478
QY 4375 ATACCACCATGATGATCGAGGTCAAGAGCATTTGCCATGATGCGAGGAGGTCTTCCA 4434
DB 3477 ATACCACCATGATGATCGAGGTCAAGAGCATTTGCCATGATGCGAGGAGGTCTTCCA 3418
QY 4435 TCTGAACTGGAGGCTTGTAAATAATTTCTTCACTTCTGCTACGACAGCAGATGAGTCTT 4494
DB 3417 TCTGAACTGGAGGCTTGTAAATAATTTCTTCACTTCTGCTACGACAGCAGATGAGTCTT 3358
QY 4495 ACGGAAATAATGAATTCATTCATTCAGATGCTCTCTGTTTATAGAAAGTACAGAAAC 4554
DB 3357 ACGGAAATAATGAATTCATTCATTCAGATGCTCTCTGTTTATAGAAAGTACAGAAAC 3298
QY 4555 ACTGACTACAGGAAACATACAGATGATAAATCTGCTTTATTTGCTGTATTGAAATCTG 4614
DB 3297 ACTGACTACAGGAAACATACAGATGATAAATCTGCTTTATTTGCTGTATTGAAATCTG 3238
QY 4615 GGTTCAGAGCCAGCAACCAACTGATGTCATATGATGATGATGATGATGATGATGATGAT 4674
DB 3237 GGTTCAGAGCCAGCAACCAACTGATGTCATATGATGATGATGATGATGATGATGAT 3178
QY 4675 CCATATCTGATTTGGTGTGAGCAGACTATCCATATCTTTAGAGTGTGTTCCCTCTAAG 4734
DB 3177 CCATATCTGATTTGGTGTGAGCAGACTATCCATATCTTTAGAGTGTGTTCCCTCTAAG 3118
QY 4735 TTTTCGGGATTTCTGGATGAGTGTGATGGGTGTCTCAAGAGGCTCAGTAGCAGCACC 4794
DB 3117 TTTTCGGGATTTCTGGATGAGTGTGATGGGTGTCTCAAGAGGCTCAGTAGCAGCACC 3058
QY 4795 AAAAGCAGAAATTTGTTACTAGGTGTTAAACAAATTCAGATGATACCAATGTCGGT 4854
DB 3057 AAAAGCAGAAATTTGTTACTAGGTGTTAAACAAATTCAGATGATACCAATGTCGGT 2998
QY 4855 ATACTCTCTCCAAAGAAAGTTCTAAGATTAGTTCATTAATTTGAGTAGTTCACTCACT 4914
DB 2997 ATACTCTCTCCAAAGAAAGTTCTAAGATTAGTTCATTAATTTGAGTAGTTCACTCACT 2938
QY 4915 AAATGTCTAGAGCTGGGCTTTTAAACAAATTAAGGAGAAAGTATCTCTCAAACTTTGATGAC 4974
DB 2937 AAATGTCTAGAGCTGGGCTTTTAAACAAATTAAGGAGAAAGTATCTCTCAAACTTTGATGAC 2878
QY 4975 ATATGCTTTACTCTGAGGTTTCCATTTTGTGTGTCACACTGCACACTCAGATCTCCGCT 5034
DB 2877 ATATGCTTTACTCTGAGGTTTCCATTTTGTGTGTCACACTGCACACTCAGATCTCCGCT 2818
QY 5035 CCGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCATGAAGAAACA 5094
DB 2817 CCGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCATGAAGAAACA 2758
QY 5095 GAGGCTGTGTGGCAACACCAACCAAGCAACCTATAGTTGATATCATCTGCTGAATCTCTGA 5154
DB 2757 GAGGCTGTGTGGCAACACCAACCAAGCAACCTATAGTTGATATCATCTGCTGAATCTCTGA 2698

PF 26-APR-2001; 2001CN-00112725.
XX
PR 26-APR-2001; 2001CN-00112725.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2003-269462/27.
XX
DR P-PSDB; ADC49296.
XX
PT Polypeptide-cell regulation factor-36.19 and polynucleotide for coding
XX
PT it.
XX
PS Claim 6; Page 25-26; 31pp; Chinese.
XX
XX The present invention discloses a novel polypeptide-cell regulation
CC factor -36.19, polynucleotide coding for the polypeptide and method for
CC producing this polypeptide by using DNA recombination technology. The
CC invention also discloses the method for curing several diseases, such as
CC cancer, malignant tumour, haemopathy, human immunodeficiency virus (HIV)
CC infection, immunological diseases and various inflammations by using the
CC said polypeptide. The invention also discloses an antagonist for resisting
CC said polypeptide and its therapeutic action and also discloses the
CC application of the polynucleotide for coding this novel Polypeptide-cell
CC regulation factor -36.19. The sequence presented is the polypeptide-cell
CC regulation factor -36.19 cDNA.
XX
SQ Sequence 1334 BP; 428 A; 254 C; 300 G; 352 T; 0 U; 0 Other;

Query Match 17.3%; Score 976.8; DB 10; Length 1334;
Best Local Similarity 97.4%; Pred. No. 2.1e-253;
Matches 993; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 19 CCCGTCATATGGCGCGATCGCGCGCGCTCTCTGAGAACCTCCGAGTACGAGG 78
DB 2 CCCGTCATATGGCGCGATCGCGCGCGCTCTCTGAGAACCTCCGAGTACGAGG 61

QY 79 CGGAATGACGCGCGAGGAGAACTCCGCTGATCTGACCGAGAACCTCTCTGATAAC 138
DB 62 CGGAATGACGCGCGAGGAGAACTCCGCTGATCTGACCGAGAACCTCTCTGATAAC 121

QY 139 TTAAGAGAGATTCCTCAAAATGTGCCAGATGTCAGGAGTATCAAAATATGAGAAAGCTA 198
DB 122 TTAAGAGAGATTCCTCAAAATGTGCCAGATGTCAGGAGTATCAAAATATGAGAAAGCTA 181

QY 199 GGCCATCTGATATCTTACTAAGCTTTTGTGATATGCGCAGTGAAGAAACTG 258
DB 182 GGCCATCTGATATCTTACTAAGCTTTTGTGATATGCGCAGTGAAGAAACTG 241

QY 259 GGCTTTTCACTATGAGGATATCATATTTGTTGGGTTAGCTTTTAAATGAAGCAAAA 318
DB 242 GGCTTTTCACTATGAGGATATCATATTTGTTGGGTTAGCTTTTAAATGAAGCAAAA 301

QY 319 GAAGTGGCAGCAGCGGCTACGAGCGCTTCGATATCTCATCCAAAGCTCCAGTATCTC 378
DB 302 GAAGTGGCAGCAGCGGCTACGAGCGCTTCGATATCTCATCCAAAGCTCCAGTATCTC 361

QY 379 CAGAGGTGCTTAAATTTGAAGTGGACTTTTATATAGCTAGGTGATGATACATACACAG 438
DB 362 CAGAGGTGCTTAAATTTGAAGTGGACTTTTATATAGCTAGGTGATGATACATACACAG 421

QY 439 AGCAACGAGGTAGAGAGACACAGCACTTCGATTAGTTCAGAAAGATGATTCTGTGAAT 498
DB 422 AGCAACGAGGTAGAGAGACACAGCACTTCGATTAGTTCAGAAAGATGATTCTGTGAAT 481

QY 499 GCTTCCTGTTTCTAGTCTGTGACCAACTCATTAATTCAGTTGGAATGATGAGACTT 558
DB 482 GCTTCCTGTTTCTAGTCTGTGACCAACTCATTAATTCAGTTGGAATGATGAGACTT 541

QY 559 CAAGAAAGACAGAAATGGTCCGAGCATGCTTCCATTTATCTGTAAGTACGACTTCAG 618
DB 542 CAAGAAAGACAGAAATGGTCCGAGCATGCTTCCATTTATCTGTGAAGTACGACTTCAG 601

QY 619 AATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAAAATGTGATTGAT 678
DB 602 AATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAAAATGTGATTGAT 661

QY 679 TGCCAAATTAAGTCCGAATAAATGAGGCGCTTAATTAACAATTTTGGACCTTCTTAATCAT 738
DB 662 TGCCAAATTAAGTCCGAATAAATGAGGCGCTTAATTAACAATTTTGGACCTTCTTAATCAT 721

QY 739 CCAAGAGCTCGGAGTATGTCGAGCTGATAGTAAATTTAGAGAGAAATTTTAGCACCTTAT 798
DB 722 CCAAGAGCTCGGAGTATGTCGAGCTGATAGTAAATTTAGAGAGAAATTTTAGCACCTTAT 781

QY 799 ACTGATTTTCACTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAAAGAGACAGA 858
DB 782 ACTGATTTTCACTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAAAGAGACAGA 841

QY 859 GAAGCAGATTTTTCAGCAGTAAATGGGAATCATAGCAACATTCGAGTATGGCAGGT 918
DB 842 GAAGCAGATTTTTCAGCAGTAAATGGGAATCATAGCAACATTCGAGTATGGCAGGT 901

QY 919 ATTATTAAATTTATGTAACCTGGAAATCTGGGATCCAGTCTCTAATAGGAGTACTTTGC 978
DB 902 ATTATTAAATTTATGTAACCTGGAAATCTGGGATCCAGTCTCTAATAGGAGTACTTTGC 961

QY 979 ATACCAATATGGAATAAAGCGAGGTCTACTTGAAGTGTCTTATGATATATTTTCGTCTT 1038
DB 962 ATACCAATATGGAATAAAGCGAGGTCTACTTGAAGTGTCTTATGATATATTTTCGTCTT 1021

RESULT 10
ADRO8103
ID ADRO8103 standard; cDNA; 3129 BP.
XX
AC ADRO8103;
XX
DT 04-NOV-2004 (first entry)
XX
DE Full length human cDNA useful for treating neurological disease Seq 1609.
XX
KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
XX
PR 09-MAY-2003; 2003JP-00131452.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
XX Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI; 2004-583265/57.
XX
XX P-PSDB; ADRI0059.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 1609; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
XX molecules and the encoded proteins thereof. Specifically, it refers to

cDNA clones obtained by an oligo-capping method, where none of these clones are identical to any known human mRNAs. The present invention describes an immunoassay to identify agonists and antagonists, as well as antibodies, antisense molecules and siRNAs that can all be used to bind to and modulate expression of the cDNA molecules. As such, these molecules are useful for diagnostic markers or therapeutic targets for the various diseases or morbid states. In particular, they are useful in gene therapy for treating osteoporosis, neurological disease, Alzheimer's disease, Parkinson's disease, dementia, short memory and various cancers, as well as for maintaining equilibrium of sense or motor function, and for treating emotional reaction, fear response and panic. Accordingly, they exhibit osteopathic, neuroprotective, neurotropic, antiparkinsonian, cyostatic and tranquiliser activities. This polynucleotide is a full length human cDNA sequence of the invention. NOTE: This sequence is not given in the sequence listing of the specification but can be obtained on CD-ROM from the European Patent Office, Vienna Sub-office.

SQ Sequence 3129 BP; 1002 A; 563 C; 694 G; 870 T; 0 U; 0 Other;

```
Query Match      15.5%; Score 874.8; DB 13; Length 3129;
Best Local Similarity 97.1%; Pred. No. 1.4e-225;
Matches 891; Conservative 0; Mismatches 27; Indels 0;
```

QY	121	CGAGAACCTTCTGATAA	CTTAAAGAGAGATTCTCAAAATGTGCCAGATTGCAGGGAGTA	180
DB	58	CGAGAACCTTCTGATAA	CTTAAAGAGAGATTCTCAAAATGTGCCAGATTGCAGGGAGTA	117
QY	181	TCAAATATGAGAAAGCTAGGCCATCTCGAATAA	CTTTTACTAAGCTTCCTTGTGATATTGGC	240
DB	118	TCAAATATGAGAAAGCTAGGCCATCTCGAATAA	CTTTTACTAAGCTTCCTTGTGATATTGGC	177
QY	241	CACAGTGAAGAAAACTGGGCTTTTCACTATGAGGATATCATAA	TTTGTTCGGTTAGCT	300
DB	178	CACAGTGAAGAAAAATTTGGGCTTTTCACTATGAGGATATCATAA	TTTGTTCGGTTAGCT	237
QY	301	TTATTAATATGAACAAAGAAGTCGAGCAGCAGAGGCTACGAGCGCTTCGATATCTCATC	360	
DB	238	TTATTAATATGAACAAAGAAGTCGAGCAGCAGAGGCTACGAGCGCTTCGATATCTCATC	297	
QY	361	CAAGACTCCAGTATTTCTCCAGAAGTGCTAAAAATTCGAAAGTGACATATTTAATACCTAGG	420	
DB	298	CAAGACTCCAGTATTTCTCCAGAAGTGCTAAAAATTCGAAAGTGACATATTTAATACCTAGG	357	
QY	421	TGCATTGACATACAACAGAGCAAGAGGTAGAGGACACAAGCACTTCGATTAGTCAGA	480	
DB	358	TGCATTGACATACAACAGAGCAACAGAGGTAGAGGACACAAGCACTTCGATTAGTCAGA	417	
QY	481	AAGATGATTA	CTGTGTAATGCTTCTCTGTTCCTAGTTCTGTGACCAATCTCATTAATGCA	540
DB	418	AAGATGATTA	CTGTGTAATGCTTCTCTGTTCCTAGTTCTGTGACCAATCTCATTAATGCA	477
QY	541	GTGGAAATGATGGA	CTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATTCGCATTATC	600
DB	478	GTGGAAATGATGGA	CTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATTCGCATTATC	537
QY	601	TGTGAACTTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGAGGAGCATAAACACATA	660	
DB	538	TGTGAACTTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGAGGTGGCATAAACACATC	597	
QY	661	TTGAAAAATGTGATTGATTGCCAAATTAAGTCGAATAAATGAGCGCCCTAATTA	TACTACAAAT	720
DB	598	TTGAAAAATGTGATCGATTGCCAAATTAAGTCGNAATAAATGAGCGCCCTAATTA	TACTACAAAT	657
QY	721	TTGCACCTTTTAAATCATCAAAAGACTGGCAGTATGTGCGAGCTGATGTAGAATTAGAG	780	
DB	658	TTGCACCTTTTAAATCATCAAAAGACTGCACAGTATGTGCGAGCTGATGTAGAATTAGAG	717	
QY	781	AGAAATTTAGCACCTTATCTGATTTTCTACTACAGACATAGTCCAGATACAGCTTGAGGA	840	
DB	718	AGAAATTTAGCACCTTATCTGATTTTCTACTACAGACATAGTCCAGATACAGCTTGAGGA	777	
QY	841	CAGCTCAAGAAAGACAGAGAAGCACGATTTCTAGCCAGTAAATGGGAATCATAGCAACA	900	


```
CC CD-ROM from the European Patent Office, Vienna Sub-office.
XX Sequence 4726 BP; 1525 A; 730 C; 788 G; 1683 T; 0 U; 0 Other;
SQ

Query Match      14.3%; Score 809.6; DB 13; Length 4726;
Best Local Similarity 99.2%; Pred. No. 8.2e-208;
Matches 845; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

Qy 4812 ACTAGGTGTTAAACAATTCAGATGATACACCAATGTGCGGTATATCTCTTCGCAAGA 4871
Db 3 ACTAGGTGTTAAACAATTCAGATGATACACCAATGTGCGGTATATCTCTTCGCAAGA 62

Qy 4872 AGTTCTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAACTAAATGTCAAGACTGG 4931
Db 63 AGTTCTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAACTAAATGTCAAGACTGG 122

Qy 4932 GCTTTTAAACAATTAAGGAGTAGTTCCTCAACAATTTTGATGACATATGCCCTTACTCTGA 4991
Db 123 GCTTTTAAACAATTAAGGAGTAGTTCCTCAACAATTTTGATGACATATGCCCTTACTCTGA 182

Qy 4992 GGTTCCTCCATTTGCTGCACACTGCACATTCAGACTTCCGTTGCGAGGTTTCATACAAGA 5051
Db 183 GGTTCCTCCATTTGCTGCACACTGCACATTCAGACTTCCGTTGCGAGGTTTCATACAAGA 242

Qy 5052 ATTATTTCAAGATTACAGTTTCTCAAAATGCATGAAGAAGCAGAGGCTGTTGGCAAC 5111
Db 243 ATTATTTCAAGATTACAGTTTCTCAAAATGCATGAAGAAGCAGAGGCTGTTGGCAAT 302

Qy 5112 ACCACCAAGCAACTATAGTTGATACATCTGCTGAATCCTGACCTCATATTTATGATGG 5171
Db 303 ACCACCAAGCAACTATAGTTGATACATCTGCTGAATCCTGACCTCATATTTATGATGG 362

Qy 5172 ATATAGATACATATATATATTTATATTTGTCGATTTCTTAAAGCCTCAGAAAATAC 5231
Db 363 ATATAGATACATATATATATTTATATTTGTCGATTTCTTAAAGCCTCAGAAAATAC 422

Qy 5232 GACTGACTAGGACGAAAGACAGGAGTATCTTCTGACACTGTTCCGAGTTACTGTTAC 5291
Db 423 GACTGACTAGGACGAAAGACAGGAGTATCTTCTGACACTGTTCCGAGTTACTGTTAC 482

Qy 5292 ATGAACAGTTGGAACTGCTGACTTCTTAAACCAAAACAACTTCTCTCTCTCTTTGTA 5351
Db 483 ATGAACAGTTGGAACTGCTGACTTCTTAAACCAAAACAACTTCTCTCTCTCTTTGTA 542

Qy 5352 GCCTTTTGAGGGTTTCATGATTTCAATTAACCAAGTTTAAAGAGTTTCAGTTACCAATGTAT 5411
Db 543 GCCTTTTGAGGGTTTCATGATTTCAATTAACCAAGTTTAAAGAGTTTCAGTTACCAATGTAT 602

Qy 5412 GCAAGAGCCAGCACTGAATACCTACATAGGTTTCTATTTCTTCAATTTTAAAGCGT 5471
Db 603 GCAAGAGCCAGCACTGAATACCTACATAGGTTTCTATTTCTTCAATTTTAAAGCGAT 662

Qy 5472 AATGACAGTGAACAATAATGGGATATGCGAAGACCCCTTCAACAAGTTATTTCTGAATG 5531
Db 663 AATGACAGTGAACAATAATGGGATATGCGAAGACCCCTTCAACAAGTTATTTCTGAATG 722

Qy 5532 A-TTTTGGGTAAATAATACAGATGCTTGTATGTTAACTAACTTGTGGAAGCAGGAAT 5590
Db 723 ATTTTGGGTAAATAATACAGATGCTTGTATGTTAACTAACTTGTGGAAGCAGGAAT 782

Qy 5591 CAGTGTCTTAAGGCTGCATCTTATTAACCAAAATGGGGTGTGTGTATTAACCTGGCTGGTAT 5650
Db 783 CAGTGTCTTAAGGCTGCATCTTATTAACCAAAATGGGGTGTGTGTATTAACCTGGCTGGTAT 840

Qy 5651 TAGAGAGGGAAC 5662
Db 841 TAGAGAGGGAAC 852

RESULT 12
AAS65246
ID AAS65246 standard; cDNA; 1292 BP.
XX
```

```
AC AAS65246;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #1050.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG01059.
XX CC New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 1050; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1292 BP; 365 A; 261 C; 312 G; 354 T; 0 U; 0 Other;

Query Match      13.1%; Score 744; DB 5; Length 1292;
Best Local Similarity 88.4%; Pred. No. 2.4e-190;
Matches 879; Conservative 0; Mismatches 0; Indels 115; Gaps 2;

Qy 4162 AGGTTTCATGAAGCCCTTAAGTTATGTCATCATTAAGATAAAGAGATTTTATTGAGTCCCTATT 4221
Db 413 AGGTTTCATGAAGCCCTTAAGTTATGTCATCATTAAGATAAAGAGATTTTATTGAGTCCCTATT 472

Qy 4222 AATCAAAATACCTGCAACGATCTTCTCAGTGGGTCCATGTGTGTCAGTGCCACATAT 4281
Db 473 AATCAAAATACCTGCAACGATCTTCTCAGTGGGTCCATGTGTGTCAGTGCCACATAT 532

Qy 4282 GGGGGTTTCAGATGATTACATTTGCTTTGCTCTCCCGGTGGATATAAATGATATATTTCCAG 4341
Db 533 GGGGGTTTCAGATGATTACATTTGCTTTGCTCTCCCGGTGGATATAAATGATATATTTCCAG 592
```


Qy 720 TTTCACCTTCTTAATCATCCAAAGACTCGCAGTATGTGCGAGCTGATGTAGAAATTAGA 779
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 670 TTTCACCTTCTTAATCATCCAAAGACTC-GCAGTATGTGCGCACTGATGTAGAAANTAGA 728
 Qy 780 GAGAAATTTAGCACC 794
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 729 GAGAAATTTAGCACC 743

RESULT 14

AAS65245
 ID AAS65245 standard; cDNA; 1185 BP.

XX AC
 XX AAS65245;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #1049.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG01058.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 1049; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridization probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1185 BP; 360 A; 262 C; 270 G; 293 T; 0 U; 0 Other;

XX Query Match

9.5%; Score 537.8; DB 5; Length 1185;

Best Local Similarity 85.1%; Pred. No. 1.7e-134;
 Matches 663; Conservative 0; Mismatches 2; Indels 114; Gaps 1;

Qy 4163 GGTTTCATGAAGCCTTAAGTTATGTCATCATTAGATAAAGAGATTTTATTCAGTCCCTATTA 4222
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 302 GGTTTCATGAAGCCTTAAGTTATGTCATCATTAGATAAAGAGATTTTATTCAGTCCCTATTA 361
 Qy 4223 ATCAAAATACCCCTGCAACGATCTTCCTCAGTGGGGTCCATGGTGTCCAGTGCCACATATG 4282
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 362 ATCAAAATACCCCTGCAACGATCTTCCTCAGTGGGGTCCATGGTGTCCAGTGCCACATATG 421
 Qy 4283 GGGGTTTCAGATGATTAATTCGTTCTCTCCGGTGGATATAAATGATATATTCAGG 4342
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 422 GGGGTTTCAGATGATTAATTCGTTCTCTCCGGTGGATATAAATGATATATTCAGG 481
 Qy 4343 TAAAGGATATTCCTCTATTTTCAGACAAAAAATACCAACACATGATGATCGAGGTCAA 4402
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 482 TAAAGGATATTCCTCTATTTTCAGACAAAAAATACCAACACATGATGATCGAGGTCAA 541
 Qy 4403 GAGCATTTGCCATGATGACAGGAGGTCTTCCATCTGGAACTGGAGGTCTTTGTAATAAATT 4462
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 542 GAGCATTTGCCATGATGACAGG----- 563
 Qy 4463 CTTTTCATCTGCTACGACACAGATGAGTCTTACGGAATAATGAATTCATTCATTCAG 4522
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 564 ----- 563
 Qy 4523 ATGCCTCTCTGTTTTTAGAAAGTACAGAGACACTGGACTACAGGACATACAGATGATA 4582
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 564 -----AGAAAGTACAGAGACACTGGACTACAGGAAACATACAGATGATA 607
 Qy 4583 ACTGCCCTTTATTTGTTCTGTATTTGAAATCTGGGTTTCCAGCCAGCAACCAACTGAGTG 4642
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 608 ACTGCCCTTTATTTGTTCTGTATTTGACATTTCTGGTTTCCAGCCAGCAACCAACTGAGTG 667
 Qy 4643 CAATATGTAGTCATTCAGACTTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAGCTA 4702
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 668 CGATATGTAGTCATTCAGACTTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAGCTA 727
 Qy 4703 TCCATAATCCTTTAGAGTGGTTCCCTCTAAGTTTTCGGGGATTTCTGGATGCGAGTCATG 4762
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 728 TCCATAATCCTTTAGAGTGGTTCCCTCTAAGTTTTCGGGGATTTCTGGATGCGAGTCATG 787
 Qy 4763 GGGTGTCTCAAGAGGCTCAGCTAGCAGCACCAAAAGCAGAGAAATTTCTTACTAGGTGTTA 4822
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 788 GGGTGTCTCAAGAGGCTCAGCTAGCAGCACCAAAAGCAGAGAAATTTCTTACTAGGTGTTA 847
 Qy 4823 AAACAATTCAGATGATATACACCAATGTGCGGTATATCTCTTCGCAAGAAAGTTCTTAAGAT 4882
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 848 AAACAATTCAGATGATATACACCAATGTGCGGTATATCTCTTCGCAAGAAAGTTCTTAAGAT 907
 Qy 4883 TAGTCATTAATTTAGTGTAGTTTCAAGTTTCAACTAAATGTCATGAGCTGGGCTTTTAAACA 4941
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 908 TAGTCATTAATTTAGTGTAGTTTCAAGTTTCAACTAAATGTCATGAGCTGGGCTTTTAAACA 966

RESULT 15

AAS80497
 ID AAS80497 standard; cDNA; 1185 BP.

XX AC
 XX AAS80497;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #16301.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG16310.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 16301; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Search completed: April 8, 2005, 19:02:03
Job time : 1901 secs

Query Match 9.5%; Score 537.8; DB 5; Length 1185;
Best Local Similarity 85.1%; Pred. No. 1.7e-134;
Matches 663; Conservative 0; Mismatches 2; Indels 114; Gaps 1;
QY 4163 GGTTCATGAAGCCCTTAAGTTATGTCATCATTAAGATAAAGAGATTTATTGAGTCCCTATTA 4222
DB 302 GGTTCATGAAGCCCTTAAGTTATGTCATCATTAAGATAAAGAGATTTATTGAGTCCCTATTA 361
QY 4223 ATCAAAATACCTCGCAACGATCTTCTCAGTGGGTCCAGTGTCGACCATATG 4282
DB 362 ATCAAAATACCTCGCAACGATCTTCTCAGTGGGTCCAGTGTCGACCATATG 421
QY 4283 GGGGTTTCAGATGATACATTTGGTCTTCTCCGGTGGATATAATGATATATTCCAGG 4342
DB 422 GGGGTTTCAGATGATACATTTGGTCTTCTCCGGTGGATATAATGATATATTCCAGG 481
QY 4343 TAAAGGATATTCCTTATTTTTCAGCAAAAAAATACACCATCATGATCGAGGTGCAA 4402
DB 482 TAAAGGATATTCCTTATTTTTCAGCAAAAAAATACACCATCATGATCGAGGTGCAA 541
QY 4403 GAGCATTTGCCCATGATGAGAGGTCTTCCATCTGGAACTGGAGGTCTTGTAAAAAAT 4462
DB 542 GAGCATTTGCCCATGATGAGG----- 563
QY 4463 CTTTTTCACCTGTACGACGACGATGAGTCTTACGGAAATATGAATTCATTCATTCCAG 4522
DB 564 ----- 563

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 14:46:37 ; Search time 15970.4 Seconds
(without alignments)
17178.818 Million cell updates/sec

Title: US-10-782-244-1
Perfect score: 5662
Sequence: 1 agcgggtgtgactgaacc.....gctggtattagagagggaac 5662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5662	100.0	5662	6	CQ869395
2	5662	100.0	5662	9	AY515854 Homo sapi
3	5127	90.6	5127	6	CQ869396 Sequence
4	4363.4	77.1	9297	10	AY540053 Mus muscu
5	4300.4	76.0	8213	9	AB082530 Homo sapi
6	4271.4	75.4	8232	6	BD185164 Novel gen
7	4158.8	73.5	5127	10	AY497009 Mus muscu
8	3829	67.6	3829	6	CQ847976 Sequence
9	3622.6	64.0	5440	9	BC051729 Homo sapi
10	2244.2	39.6	4657	6	CQ411997 Sequence
11	2239.4	39.6	4601	6	CQ488314 Sequence
12	2239.4	39.6	4602	6	CQ494141 Sequence
13	2220.8	39.2	6816	10	AK173326 Mus muscu
14	1992.6	35.2	2430	10	BC022906 Mus muscu
15	1494	26.4	1649	6	CQ726625 Sequence
16	1009.4	17.8	97240	9	AC109467 Homo sapi
17	1009.4	17.8	170996	2	AC010938 Homo sapi
18	1009.4	17.8	198748	2	AC016719 Homo sapi
19	874.8	15.5	3129	6	CQ851140 Sequence

20	874.8	15.5	3129	9	AK128358 Homo sapi
21	830.2	14.7	214225	2	AC105581 Rattus no
22	809.6	14.3	4726	6	CQ850194 Sequence
23	809.6	14.3	4726	9	AK127305 Homo sapi
24	768.6	13.6	1778	9	BC029608 Homo sapi
25	763	13.5	153145	10	AC102825 Mus muscu
26	680.2	12.0	4851	10	BC058643 Mus muscu
27	612.8	10.8	765	6	BD020210 Novel gen
28	612.8	10.8	765	6	BD100148 Novel gen
29	534	9.4	534	6	CQ726624 Sequence
30	515.2	9.1	68372	2	AC023763 Homo sapi
31	465.2	8.2	570	6	CQ392940 Sequence
32	465.2	8.2	570	6	CQ399338 Sequence
33	392.4	6.9	401	6	AX885079 Sequence
34	392.4	6.9	401	6	BD024689 Sequence
35	385.6	6.8	452	6	CQ477511 Sequence
36	383.4	6.8	427	6	CQ498714 Sequence
37	383.4	6.8	427	6	CQ507684 Sequence
38	371.6	6.6	973	5	CR391153 Gallus ga
39	371	6.6	772	6	BD020211 Novel gen
40	371	6.6	772	6	BD100149 Novel gen
41	336.4	5.9	416	6	CQ468343 Sequence
42	334	5.9	341	6	CQ405723 Sequence
43	264.6	4.7	164698	2	EX936441 Danio rer
44	264.6	4.7	186000	2	CR735111 Danio rer
45	196	3.5	168788	2	CR318610 Danio rer

ALIGNMENTS

RESULT 1
LOCUS CQ869395 5662 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 1 from Patent WO2004074448.
ACCESSION CQ869395
VERSION CQ869395.1 GI:51999314
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Sabatini,D.M., Kim,D.H. and Sarbassov,D.D.
TITLE ~~PI3K-KINASE-ASSOCIATED PROTEINS~~
JOURNAL Patent: WO 2004074448-A 1 02-SEP-2004;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES
Location/Qualifiers
1..5662
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN	Query Match	100.0%;	Score 5662;	DB 6;	Length 5662;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 5662;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	AGCGGGTGTGACTGAAGAACCCGTCATATATGCGCGGATCGGCGCGCGCTCTCTGAAG	60		
Db	1	AGCGGGTGTGACTGAAGAACCCGTCATATGCGCGGATCGGCGCGCGCTCTCTGAAG	60		
Qy	61	AACCTCCGAGTACGAGGGCGGAATGACAGCGCGGAGAGACGTCGCCGCTGATCTGACC	120		
Db	61	AACCTCCGAGTACGAGGGCGGAATGACAGCGCGGAGAGACGTCGCCGCTGATCTGACC	120		
Qy	121	CGAGAACCTTCTGATACCTTAAGAGAGATTCTCCAAAATGTGGCCAGATTGCAGGGAGTA	180		
Db	121	CGAGAACCTTCTGATACCTTAAGAGAGATTCTCCAAAATGTGGCCAGATTGCAGGGAGTA	180		
Qy	181	TCAAATATGAGAAAGCTAGGCCATCTGAATACTTTTACTAAGCTTCTTTGTGATATTGGC	240		
Db	181	TCAAATATGAGAAAGCTAGGCCATCTGAATACTTTTACTAAGCTTCTTTGTGATATTGGC	240		

QY	241	CACAGTGAAGAAAACTGGGCTTTTCACTATGAGGATATCATAAATTTGTTTGGCGGTAGCT	300
Db	241	CACAGTGAAGAAAACTGGGCTTTTCACTATGAGGATATCATAAATTTGTTTGGCGGTAGCT	300
QY	301	TTATTAAATGAAGCAAAAGAGTGGAGCAGCAGGCTACGAGCGCTTCGATATCTCATC	360
Db	301	TTATTAAATGAAGCAAAAGAGTGGAGCAGCAGGCTACGAGCGCTTCGATATCTCATC	360
QY	361	CAAGACTCCAGTATTTCTCCAGAAAGTGTCTAAATTTGAAAGTGGACTATTTAATAGTAGG	420
Db	361	CAAGACTCCAGTATTTCTCCAGAAAGTGTCTAAATTTGAAAGTGGACTATTTAATAGTAGG	420
QY	421	TGCATTGACATACAAAGAGCAACGAGGTAGAGAGGACAAAGCACTTCGATAGTAGCAGA	480
Db	421	TGCATTGACATACAAAGAGCAACGAGGTAGAGAGGACAAAGCACTTCGATAGTAGCAGA	480
QY	481	AAAGATGATTACTGTGATGCTTCTCTGTTTCTCTGTTTCTGAGCACTTCATTAATTGCA	540
Db	481	AAAGATGATTACTGTGATGCTTCTCTGTTTCTCTGTTTCTGAGCACTTCATTAATTGCA	540
QY	541	GTGTGAAATGATGGACTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATTGCCAATTATC	600
Db	541	GTGTGAAATGATGGACTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATTGCCAATTATC	600
QY	601	TGTGAATAGCACTTCAGAAATCCAGAGGTGGTGGCTTCGAGGAGGACTAAACACCATTA	660
Db	601	TGTGAATAGCACTTCAGAAATCCAGAGGTGGTGGCTTCGAGGAGGACTAAACACCATTA	660
QY	661	TTGAAATGTCGATGTCGATTTGCCAATTTAAGTCCGAATAAATGAGGCCCTAATTAACAATT	720
Db	661	TTGAAATGTCGATGTCGATTTGCCAATTTAAGTCCGAATAAATGAGGCCCTAATTAACAATT	720
QY	721	TTGCACCTTCTTAATCATCAAGACTCCGAGCTATGTCGAGCTGATGTAGAAATTAGAG	780
Db	721	TTGCACCTTCTTAATCATCAAGACTCCGAGCTATGTCGAGCTGATGTAGAAATTAGAG	780
QY	781	AGAAATTTAGCACCTTATAGTATTTCTACTACAGACATAGTCCAGATACAGCTGAAGGA	840
Db	781	AGAAATTTAGCACCTTATAGTATTTCTACTACAGACATAGTCCAGATACAGCTGAAGGA	840
QY	841	CAGCTCAAGAAAGACAGAGAACGAGATTTCTAGCCAGTAAATGGGAATCATAGCAACA	900
Db	841	CAGCTCAAGAAAGACAGAGAACGAGATTTCTAGCCAGTAAATGGGAATCATAGCAACA	900
QY	901	TTCCGATCATGGGCAGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	960
Db	901	TTCCGATCATGGGCAGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	960
QY	961	CTAATAGGAGTACTTTTGCATACCAAAATATGGAATAAGGCGAGGTCTACTTGAAGTGCTT	1020
Db	961	CTAATAGGAGTACTTTTGCATACCAAAATATGGAATAAGGCGAGGTCTACTTGAAGTGCTT	1020
QY	1021	TATGATATATTTGCTCTTCTCTCTACCTGTTGTGATGAGAGTTCTATAGAACCACTACTC	1080
Db	1021	TATGATATATTTGCTCTTCTCTCTACCTGTTGTGATGAGAGTTCTATAGAACCACTACTC	1080
QY	1081	AGTGTAGATCCAGGGAGTTTCCAAAGACAGTTGGAGGCTTTCAGATGGCTTTTGGCAGCT	1140
Db	1081	AGTGTAGATCCAGGGAGTTTCCAAAGACAGTTGGAGGCTTTCAGATGGCTTTTGGCAGCT	1140
QY	1141	GAGGCAAAACATAATTTCTCTCATCGTGCCAGATCCAGGCCAGACCTCATGGATAATTAT	1200
Db	1141	GAGGCAAAACATAATTTCTCTCATCGTGCCAGATCCAGGCCAGACCTCATGGATAATTAT	1200
QY	1201	TTGGCAGCTGATCTCTGCAATTTATTCGTAATGGAATTTTAGGGGTCTAGTTGAAGTG	1260
Db	1201	TTGGCAGCTGATCTCTGCAATTTATTCGTAATGGAATTTTAGGGGTCTAGTTGAAGTG	1260
QY	1261	ATAACAAACAGTGTATCATCTCAGTTAGACTACCATCTCTTTTAGGAGAGCTTTTA	1320
Db	1261	ATAACAAACAGTGTATCATCTCAGTTAGACTACCATCTCTTTTAGGAGAGCTTTTA	1320
QY	1321	CATATGGCAAAACAATTTCTTCTCATTTACATAGCCATCATTTACATGCTTTGCCAACC	1380
Db	1321	CATATGGCAAAACAATTTCTTCTCATTTACATAGCCATCATTTACATGCTTTGCCAACC	1380
QY	1381	CTAATGAATATGCTGCTATCTTTTGATATCCCAAGAAAAGAGACTGCGAGCCAGTGCA	1440
Db	1381	CTAATGAATATGCTGCTATCTTTTGATATCCCAAGAAAAGAGACTGCGAGCCAGTGCA	1440
QY	1441	GCCTTGAACTGTTTAAACCGCTTCCATGAATGAAGAAACGAGACCTTAAGCCCTTATAGT	1500
Db	1441	GCCTTGAACTGTTTAAACCGCTTCCATGAATGAAGAAACGAGACCTTAAGCCCTTATAGT	1500
QY	1501	CTTCATTTAGACCACTATTTTCAGAAAGCAATTTGCAACACACAGAAACGGGATCAGTAT	1560
Db	1501	CTTCATTTAGACCACTATTTTCAGAAAGCAATTTGCAACACACAGAAACGGGATCAGTAT	1560
QY	1561	CTCCGAGTTTCAGAAAGATATATTTATCTTTAAGGATACAGAGGAGCTCTTTTAATTAAC	1620
Db	1561	CTCCGAGTTTCAGAAAGATATATTTATCTTTAAGGATACAGAGGAGCTCTTTTAATTAAC	1620
QY	1621	CTTTAGAGATAGCCAAAGTCTTCAACATAAAGAGAACTTTGAAATGGAATCTTTATA	1680
Db	1621	CTTTAGAGATAGCCAAAGTCTTCAACATAAAGAGAACTTTGAAATGGAATCTTTATA	1680
QY	1681	GGGACCAATTTTAAAGTGGCCAAATGTAATCTAAGAACTATAAAGATGAACAGTTACAC	1740
Db	1681	GGGACCAATTTTAAAGTGGCCAAATGTAATCTAAGAACTATAAAGATGAACAGTTACAC	1740
QY	1741	AGGTTTGTACGAAGACTACTTTATTTTACAAGCCACAGCAGTAAATATATGCCAACCTG	1800
Db	1741	AGGTTTGTACGAAGACTACTTTATTTTACAAGCCACAGCAGTAAATATATGCCAACCTG	1800
QY	1801	GATCTGGAATTTTGGCAAGGCCAAACAGCTCAACGTTGTAGTTGCCAGTTTACAGAAATTT	1860
Db	1801	GATCTGGAATTTTGGCAAGGCCAAACAGCTCAACGTTGTAGTTGCCAGTTTACAGAAATTT	1860
QY	1861	CTTCTTGAATCTGAAGAGGATGGGCAAGGCTTCTAGAGATCTAGTAAAGATATTTGTT	1920
Db	1861	CTTCTTGAATCTGAAGAGGATGGGCAAGGCTTCTAGAGATCTAGTAAAGATATTTGTT	1920
QY	1921	CAGTGGCTCAATGCTTCTCATCTGGAATGAAACCCGAAAGAACTCTTCAAAATTAATGGTTTA	1980
Db	1921	CAGTGGCTCAATGCTTCTCATCTGGAATGAAACCCGAAAGAACTCTTCAAAATTAATGGTTTA	1980
QY	1981	TTGACCACTTCTAGTCAACACTACTTTTATTTATTTAGGAACACTTTCTTGGCACCTCAT	2040
Db	1981	TTGACCACTTCTAGTCAACACTACTTTTATTTATTTAGGAACACTTTCTTGGCACCTCAT	2040
QY	2041	GGAGTTAAATGCTGGAAATGCACTGCTTCTAGTCTCTCTTAACTTTTGTCTCTCTTG	2100
Db	2041	GGAGTTAAATGCTGGAAATGCACTGCTTCTAGTCTCTCTTAACTTTTGTCTCTCTTG	2100
QY	2101	AAAAACCAAGATCACTTGTCTAAAACTTACTGTTTCTAGCTTGGACTATAGCAGAGATGA	2160
Db	2101	AAAAACCAAGATCACTTGTCTAAAACTTACTGTTTCTAGCTTGGACTATAGCAGAGATGA	2160
QY	2161	TTGGCTAGAGTCACTCTTTTCAAAAATTTTAACTGAGCTACTGATGCCCTGCGAGACTCTAT	2220
Db	2161	TTGGCTAGAGTCACTCTTTTCAAAAATTTTAACTGAGCTACTGATGCCCTGCGAGACTCTAT	2220
QY	2221	GCNAACAAACATTTTAAAGGTTATTTTGAAGCTAATTTGAAATCTTTTAATTAATTTGGGA	2280
Db	2221	GCNAACAAACATTTTAAAGGTTATTTTGAAGCTAATTTGAAATCTTTTAATTAATTTGGGA	2280
QY	2281	ATTGAGTTTGTAGTGACCCAGCTACATGATAAAAAACAAACGATTTTCTCTGAGCTCTT	2340
Db	2281	ATTGAGTTTGTAGTGACCCAGCTACATGATAAAAAACAAACGATTTTCTCTGAGCTCTT	2340
QY	2341	GATATCTCGATGAAGCATGTGAAGCAAGGCCAATCTTTCATGCTCTCAATTCAGATGAAA	2400
Db	2341	GATATCTCGATGAAGCATGTGAAGCAAGGCCAATCTTTCATGCTCTCAATTCAGATGAAA	2400
QY	2401	CCAGCGTTATCCCACTTGGAGCAAGGGTTTCTCTCTCGTGTGAGATTTCTCTCCANTT	2460

[illegible]

Db	3481	ACTGTACAGAAAA	CATTACAATTAGAGACTTCATTATTTGCGGAATAGCACATTGGAAGAC	3541
Qy	3541	ACTGGTAGTACACCAAGCATTGGAGAAAAATGACTTAAAAATTCACCAAGAAATTTTGGTACA	3600	
Db	3541	ACTGGTAGTACACCAAGCATTGGAGAAAAATGACTTAAAAATTCACCAAGAAATTTTGGTACA	3600	
Qy	3601	GAGAAATCACAGAGAAAAATACAAGCCGAGAGAGGTTAGTAGTAGAAAAAGTTTCAACGAGCTCA	3660	
Db	3601	GAGAAATCACAGAGAAAAATACAAGCCGAGAGAGGTTAGTAGTAGAAAAAGTTTCAACGAGCTCA	3660	
Qy	3661	CATATGAAGATACGTAGTCCAAAGTTTCAATACACAGACACTACACAAAGTGGCATAAAGTTTCA	3720	
Db	3661	CATATGAAGATACGTAGTCCAAAGTTTCAATACACAGACACTACACAAAGTGGCATAAAGTTTCA	3720	
Qy	3721	ATGAGCTCAAGTCCCTCCACGAGAGACAGTAGTAGGTAGTACACTAACAATAGGACACAGAC	3780	
Db	3721	ATGAGCTCAAGTCCCTCCACGAGAGACAGTAGTAGGTAGTACACTAACAATAGGACACAGAC	3780	
Qy	3781	TGTGGAAAGCATGAGTACTGTGGTAAGTACTTAAAACTATTAAAGACAAGCCCACTATTTGACG	3840	
Db	3781	TGTGGAAAGCATGAGTACTGTGGTAAGTACTTAAAACTATTAAAGACAAGCCCACTATTTGACG	3840	
Qy	3841	CCACAGTCTAAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTCGTGCCTCCAGGT	3900	
Db	3841	CCACAGTCTAAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTCGTGCCTCCAGGT	3900	
Qy	3901	TCCTTCATACGCTTCTTAGAAGACACAGTCGCCCTTAAAGCACCCCTCTATTGCTACAAAT	3960	
Db	3901	TCCTTCATACGCTTCTTAGAAGACACAGTCGCCCTTAAAGCACCCCTCTATTGCTACAAAT	3960	
Qy	3961	AAAAGTCTACAGATTGTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGCTATGCT	4020	
Db	3961	AAAAGTCTACAGATTGTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGCTATGCT	4020	
Qy	4021	ACACTGAAAAGACTACAGCAAAAGAAATGATCCATCTCTATCTCACTCTGAAGCTTTG	4080	
Db	4021	ACACTGAAAAGACTACAGCAAAAGAAATGATCCATCTCTATCTCACTCTGAAGCTTTG	4080	
Qy	4081	GCATCTCCAGCAAAAGATGTCTATTACTGTATACCATCACCAATGAAGGCCACACAGTTT	4140	
Db	4081	GCATCTCCAGCAAAAGATGTCTATTACTGTATACCATCACCAATGAAGGCCACACAGTTT	4140	
Qy	4141	GAGTCCAGATTACACCAAGCAGTTTCATGAAAGCCTTAAGTTATGCATCATTTAGATAAA	4200	
Db	4141	GAGTCCAGATTACACCAAGCAGTTTCATGAAAGCCTTAAGTTATGCATCATTTAGATAAA	4200	
Qy	4201	GAAGATTTATGAGTCCCTATTAAATCAAAAATACCCCTGCAACGATCTTCAGTGCGGTCC	4260	
Db	4201	GAAGATTTATGAGTCCCTATTAAATCAAAAATACCCCTGCAACGATCTTCAGTGCGGTCC	4260	
Qy	4261	ATGGTGTCCAGTCCACATATGGGGTTCAGATGATTACATTTGGTCTTCTCTCCCGTG	4320	
Db	4261	ATGGTGTCCAGTCCACATATGGGGTTCAGATGATTACATTTGGTCTTCTCTCCCGTG	4320	
Qy	4321	GATATAAATGATATATTCCAGGTAAAGGATATTCCCTATTTTCAGACAAAAAACATACCA	4380	
Db	4321	GATATAAATGATATATTCCAGGTAAAGGATATTCCCTATTTTCAGACAAAAAACATACCA	4380	
Qy	4381	CCACATGATGATCGAGGTGCAAGAGCATTTGGCCCATGATGCGAGGAGTCTTCCATCTGGA	4440	
Db	4381	CCACATGATGATCGAGGTGCAAGAGCATTTGGCCCATGATGCGAGGAGTCTTCCATCTGGA	4440	
Qy	4441	ACTGGAGGTCTTGTAAAAATCTTTTCACTTGTCTACGACAGCATGAGTCTTACGGAA	4500	
Db	4441	ACTGGAGGTCTTGTAAAAATCTTTTCACTTGTCTACGACAGCATGAGTCTTACGGAA	4500	
Qy	4501	ATAATGAAATCAATCCATTCAGATGCTCTCTGTGTTTTTGAAGAATACAGAAGACACTGGA	4560	
Db	4501	ATAATGAAATCAATCCATTCAGATGCTCTCTGTGTTTTTGAAGAATACAGAAGACACTGGA	4560	
Qy	4561	CTACAGAAACATACAGATGATAACTGCCCTTTATTTGTCCTGATTGAAATTCCTGGGTTTC	4620	
Db	4561	CTACAGAAACATACAGATGATAACTGCCCTTTATTTGTCCTGATTGAAATTCCTGGGTTTC	4620	

QY	781	AGAA	TTTTAGCAC	CCCTATAT	CTGATTTT	TTTCACTAC	AGATAGT	CCAGATAC	AGCTGAAGGA	840
Db	781	AGAA	TTTTTAGCAC	CCCTATAT	CTGATTTT	TTTCACTAC	AGATAGT	CCAGATAC	AGCTGAAGGA	840
QY	841	CAGCT	CAAGAGAC	AGAGAGAC	AGATTTT	CTAGCC	AGTAAAT	TGGGAAT	CTATAGCAACA	900
Db	841	CAGCT	CAAGAGAGAC	AGAGAGAC	AGATTTT	CTAGCC	AGTAAAT	TGGGAAT	CTATAGCAACA	900
QY	901	TTCCG	ATCATGGG	CAGGTAT	TATTAAT	TTATGT	ATAACT	CTGGAAAT	CTCTGGATCCAGTCT	960
Db	901	TTCCG	ATCATGGG	CAGGTAT	TATTAAT	TTATGT	ATAACT	CTGGAAAT	CTCTGGATCCAGTCT	960
QY	961	CTAAT	PAGAGT	ACTTTT	GTGCAT	ACCAAT	ATATGGA	AAATAGG	CGAGGTCTATCTTGAAAGTGCTT	1020
Db	961	CTAAT	PAGAGT	ACTTTT	GTGCAT	ACCAAT	ATATGGA	AAATAGG	CGAGGTCTATCTTGAAAGTGCTT	1020
QY	1021	TATGAT	ATATTT	TCGTC	TTCTCT	TACCT	GTGTG	AGCTG	AGGAGTTCATAGAGCACTACTC	1080
Db	1021	TATGAT	ATATTT	TCGTC	TTCTCT	TACCT	GTGTG	AGCTG	AGGAGTTCATAGAGCACTACTC	1080
QY	1081	AGTGT	PAGAT	CCAGG	AGGTTT	CCAAGA	CAGTTT	GGAAGG	CTTTTCAGATGGCTTTTGCGCAGCT	1140
Db	1081	AGTGT	PAGAT	CCAGG	AGGTTT	CCAAGA	CAGTTT	GGAAGG	CTTTTCAGATGGCTTTTGCGCAGCT	1140
QY	1141	GAGG	CAAAA	AACTAT	TTCTTC	CTCAT	FCGTG	CCAGAT	CCAGGCCAGACCTCATGSGATAATAT	1200
Db	1141	GAGG	CAAAA	AACTAT	TTCTTC	CTCAT	FCGTG	CCAGAT	CCAGGCCAGACCTCATGSGATAATAT	1200
QY	1201	TTGGC	ACTGAT	ACTCT	CTGCAT	TATTTT	TCGTAA	TGGACT	TTTAGAGGGTCTAGTTGAAGTG	1260
Db	1201	TTGGC	ACTGAT	ACTCT	CTGCAT	TATTTT	TCGTAA	TGGACT	TTTAGAGGGTCTAGTTGAAGTG	1260
QY	1261	ATAA	CAAA	CAGT	GAATCAT	TATCT	CAGTT	PAGAGT	TACCATCTCTTTTAGGAGAGCTTTT	1320
Db	1261	ATAA	CAAA	CAGT	GAATCAT	TATCT	CAGTT	PAGAGT	TACCATCTCTTTTAGGAGAGCTTTT	1320
QY	1321	CATAT	GCCAA	ACAA	TCATTT	CTTCCT	CATT	CCATAG	CCCATCACTGCTTTGCCAAC	1380
Db	1321	CATAT	GCCAA	ACAA	TCATTT	CTTCCT	CATT	CCATAG	CCCATCACTGCTTTGCCAAC	1380
QY	1381	CTAAT	GAAAT	TATGG	CTGCAT	CTCTT	TGATAT	CCCCA	AGGAAAAGACCTGCGAGCCAGTGCA	1440
Db	1381	CTAAT	GAAAT	TATGG	CTGCAT	CTCTT	TGATAT	CCCCA	AGGAAAAGACCTGCGAGCCAGTGCA	1440
QY	1441	GCCT	TGAAC	TGTTT	TA	AAACG	CTTCC	ATGA	AAACGAGCACTTAAGCCCTATAGT	1500
Db	1441	GCCT	TGAAC	TGTTT	TA	AAACG	CTTCC	ATGA	AAACGAGCACTTAAGCCCTATAGT	1500
QY	1501	CTTC	ATTTAG	ACCA	CATTAT	TTTC	GAAAGCA	ATTTG	CAACACACGAGATCAGTAT	1560
Db	1501	CTTC	ATTTAG	ACCA	CATTAT	TTTC	GAAAGCA	ATTTG	CAACACACGAGATCAGTAT	1560
QY	1561	CTCC	GAGTT	CAGAA	AGATAT	TATTT	TCCTTA	AGGATAC	CAGAGAGCTCTTTTAATTAAC	1620
Db	1561	CTCC	GAGTT	CAGAA	AGATAT	TATTT	TCCTTA	AGGATAC	CAGAGAGCTCTTTTAATTAAC	1620
QY	1621	CTTAG	AGATAG	CCAA	AGCTT	CTCAAC	TAA	AGAGAA	TCTTGAATGGAAATTTGGAATCTTATA	1680
Db	1621	CTTAG	AGATAG	CCAA	AGCTT	CTCAAC	TAA	AGAGAA	TCTTGAATGGAAATTTGGAATCTTATA	1680
QY	1681	GGG	ACCA	TTCTT	TAAGT	GGCC	AAATGT	AAAACT	TATAAAGATGAACAGTTTACAC	1740
Db	1681	GGG	ACCA	TTCTT	TAAGT	GGCC	AAATGT	AAAACT	TATAAAGATGAACAGTTTACAC	1740
QY	1741	AGG	TTTGT	ACGA	AGCTACT	TTTAT	TTTACA	GGCC	CAGCACTAAATATATGCGAACCTG	1800
Db	1741	AGG	TTTGT	ACGA	AGCTACT	TTTAT	TTTACA	GGCC	CAGCACTAAATATATGCGAACCTG	1800
QY	1801	GATCT	GAA	TTTGG	CAAGG	CAAA	CAGCT	CA	CGGTTGTAGTTGCCAGTTTACAGAA	1860
Db	1801	GATCT	GAA	TTTGG	CAAGG	CAAA	CAGCT	CA	CGGTTGTAGTTGCCAGTTTACAGAA	1860
QY	1861	CTTCT	TTTCA	ATCTG	AAAG	AGGAT	TCGCA	AGGCT	CTTTAGAACCTCTAGTAAAGGATATTTGTT	1920

[illegible]

2941	Db	CTTGGGCTCATAGCTATAAACCAAAACAAGGCTGTGATATCTTAAATGTCAACTGGGAT	3000
3001	QY	GCTGTGAGGCATAGTCGCAACATCTGTGGCCAGTGTTCAGATGATGTGGAAACACTC	3060
3001	Db	GCTGTGAGGCATAGTCGCAACATCTGTGGCCAGTGTTCAGATGATGTGGAAACACTC	3060
3061	QY	TGTAAATGAACCTTTCACTCATATCCCAAGCACCTCTAAGTTTGAACCTCGAGTCAACCGACTCT	3120
3061	Db	TGTAAATGAACCTTTCACTCATATCCCAAGCACCTCTAAGTTTGAACCTCGAGTCAACCGACTCT	3120
3121	QY	AGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTTGGAGGATACCGG	3180
3121	Db	AGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTTGGAGGATACCGG	3180
3181	QY	TTTTGGCAGCAGCTCTACTAGTACATTTTTCTCTGATATCAATGAAGATACAGAGGCCAACA	3240
3181	Db	TTTTGGCAGCAGCTCTACTAGTACATTTTTCTCTGATATCAATGAAGATACAGAGGCCAACA	3240
3241	QY	TTTTATGACCGATCTGGACCCATAAAGGATAAAAAATTCATTCCTCTTTCTTTGCTTCTTAGT	3300
3241	Db	TTTTATGACCGATCTGGACCCATAAAGGATAAAAAATTCATTCCTCTTTCTTTGCTTCTTAGT	3300
3301	QY	AAACTTTGTGAAGATCGTATCTTAAATTCGGTTACTTTGCTCTACAAAAACAATCGTAGT	3360
3301	Db	AAACTTTGTGAAGATCGTATCTTAAATTCGGTTACTTTGCTCTACAAAAACAATCGTAGT	3360
3361	QY	AGCAGTGATCCAAAGGAGGAAATATCATCTGAAGTAAAGCAAGCACAGCGCGAATC	3420
3361	Db	AGCAGTGATCCAAAGGAGGAAATATCATCTGAAGTAAAGCAAGCACAGCGCGAATC	3420
3421	QY	AGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGTATGATTTTACACCCATATCC	3480
3421	Db	AGAACACTTACGGAGCCAGTGTGATTTTAAATCATAGTGTATGATTTTACACCCATATCC	3480
3481	QY	ACTGTACAGAAAAATTAACAATTAGAGACTTCAATTTATGCGGGAATTAAGCACTGAAGAC	3540
3481	Db	ACTGTACAGAAAAATTAACAATTAGAGACTTCAATTTATGCGGGAATTAAGCACTGAAGAC	3540
3541	QY	ACTGCTAGTACACCAAGCANTTGGAGAAAATGACTTAAATTTACCACGAATTTTGGTACA	3600
3541	Db	ACTGCTAGTACACCAAGCANTTGGAGAAAATGACTTAAATTTACCACGAATTTTGGTACA	3600
3601	QY	GAGAAATCACAGAGAAAAATACAAGCCGAGAGAGGTAGTAGTAAAAAGTTTCAACGAGCTCA	3660
3601	Db	GAGAAATCACAGAGAAAAATACAAGCCGAGAGAGGTAGTAGTAAAAAGTTTCAACGAGCTCA	3660
3661	QY	CATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATTAAGTTCA	3720
3661	Db	CATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATTAAGTTCA	3720
3721	QY	ATGAGCTCAAGTCCCTTCCAGAGAGACAGTAGTGTAGTGTCAACTATGGNACAGAC	3780
3721	Db	ATGAGCTCAAGTCCCTTCCAGAGAGACAGTAGTGTAGTGTCAACTATGGNACAGAC	3780
3781	QY	TGTGGAAAGCATGAGTACTGTGGTAAAGTACTAAAACTATTAAAGCAAGCCACTATTGACG	3840
3781	Db	TGTGGAAAGCATGAGTACTGTGGTAAAGTACTAAAACTATTAAAGCAAGCCACTATTGACG	3840
3841	QY	CCAAGTCTAACCAATCTGTCTCTCTCCAAATCAAAATTCGGTGTCCCTGTGCTCCAGGT	3900
3841	Db	CCAAGTCTAACCAATCTGTCTCTCTCCAAATCAAAATTCGGTGTCCCTGTGCTCCAGGT	3900
3901	QY	TCCTTCTCATACGCTTCTTACAGAGCACAGTCCCTTAAAGCACCCCTCTATTGCTACAAT	3960
3901	Db	TCCTTCTCATACGCTTCTTACAGAGCACAGTCCCTTAAAGCACCCCTCTATTGCTACAAT	3960
3961	QY	AAAAAGTCTAGCAGATTGAATCTTTAGTTTACAAAGTTCTAGAGATGCTTTTGGCTATGCT	4020
3961	Db	AAAAAGTCTAGCAGATTGAATCTTTAGTTTACAAAGTTCTAGAGATGCTTTTGGCTATGCT	4020
4021	QY	ACACTGAAAAAGCTACAGCAACAAAGAAATGCAATCCATCTTATCTCACTCAAGCTTTG	4080
4021	Db	ACACTGAAAAAGCTACAGCAACAAAGAAATGCAATCCATCTTATCTCACTCAAGCTTTG	4080

```
Qy 4081 GCATCTCCAGCAAAAGATGTCTATTTACTGATACCAATCACCATCAAGGCCAACAGTTTT 4140
Db 4081 GCATCTCCAGCAAAAGATGTCTATTTACTGATACCAATCACCATCAAGGCCAACAGTTTT 4140
Qy 4141 GAGTCCAGATTAAACCAAGCAGGTTTCATGAAGCCCTTAAGTTATGCAATCATTAAGATAAA 4200
Db 4141 GAGTCCAGATTAAACCAAGCAGGTTTCATGAAGCCCTTAAGTTATGCAATCATTAAGATAAA 4200
Qy 4201 GAAGATTATTTAGTCCCTATTAATCAAAATACCTGCAACCATCTTCTCAGTCGGGTCC 4260
Db 4201 GAAGATTATTTAGTCCCTATTAATCAAAATACCTGCAACCATCTTCTCAGTCGGGTCC 4260
Qy 4261 ATGTGTCCAGTCCACATATGCGGGTTTCAGATGATTAATCATTTGCTCTCCCGGTG 4320
Db 4261 ATGTGTCCAGTCCACATATGCGGGTTTCAGATGATTAATCATTTGCTCTCCCGGTG 4320
Qy 4321 GATATAATGATATATTCAGGTAAAGGATATTCCTTATTTTCAGACAAAACATACCA 4380
Db 4321 GATATAATGATATATTCAGGTAAAGGATATTCCTTATTTTCAGACAAAACATACCA 4380
Qy 4381 CCACATGATGATCAGGTGCAAGAGCAATTCGCCATGATGAGGAGTCTTCCATCTGGA 4440
Db 4381 CCACATGATGATCAGGTGCAAGAGCAATTCGCCATGATGAGGAGTCTTCCATCTGGA 4440
Qy 4441 ACTGGAGGTCCTGTAATAAAATTTCTTTTCACTTGCTACGACAGATGAGTCTTACCGAA 4500
Db 4441 ACTGGAGGTCCTGTAATAAAATTTCTTTTCACTTGCTACGACAGATGAGTCTTACCGAA 4500
Qy 4501 ATATGATTAATCAATCCATTCAGATGCTCTCTGTTTTTGAAGTACAGAGACATCGGA 4560
Db 4501 ATATGATTAATCAATCCATTCAGATGCTCTCTGTTTTTGAAGTACAGAGACATCGGA 4560
Qy 4561 CTACAGGAACATACAGATGATTAATGCTTTATTTGCTGATTTGAAATCTGGGTTTC 4620
Db 4561 CTACAGGAACATACAGATGATTAATGCTTTATTTGCTGATTTGAAATCTGGGTTTC 4620
Qy 4621 CAGCCAGCAACCACTGAGTGCAATATGATGTCATTCAGACATTTCAAGATATTTCCATAT 4680
Db 4621 CAGCCAGCAACCACTGAGTGCAATATGATGTCATTCAGACATTTCAAGATATTTCCATAT 4680
Qy 4681 TCTGATTTGGTGTAGCAGACATTCATTAATCTCTTTAGAGTGGTTCCTCTAAGTTTTTCG 4740
Db 4681 TCTGATTTGGTGTAGCAGACATTCATTAATCTCTTTAGAGTGGTTCCTCTAAGTTTTTCG 4740
Qy 4741 GGGATTTCTGATGCGAGTGGGCTCTCAGAGGCTCAGTAGCAGCACCAGAAC 4800
Db 4741 GGGATTTCTGATGCGAGTGGGCTCTCAGAGGCTCAGTAGCAGCACCAGAAC 4800
Qy 4801 ACAGAAATTTTACTAGTGTATAAAACAAATTCAGATGATACACCAATGTGCGGTACTC 4860
Db 4801 ACAGAAATTTTACTAGTGTATAAAACAAATTCAGATGATACACCAATGTGCGGTACTC 4860
Qy 4861 CTTGCAAAAGAGTTCTAAGATTAGTCAATTAATTTAGTAGTTCAGTTTCAACTAAATGT 4920
Db 4861 CTTGCAAAAGAGTTCTAAGATTAGTCAATTAATTTAGTAGTTCAGTTTCAACTAAATGT 4920
Qy 4921 CATGAGACTGGCTTTTAAACAATTAAGAGAGTATCTCTCAACATTTGATGATATGC 4980
Db 4921 CATGAGACTGGCTTTTAAACAATTAAGAGAGTATCTCTCAACATTTGATGATATGC 4980
Qy 4981 CTTTACTCTGAGGTTTCCCATTTGCTGTCACCTGACATTCAGACTTCGGTTCGGAGG 5040
Db 4981 CTTTACTCTGAGGTTTCCCATTTGCTGTCACCTGACATTCAGACTTCGGTTCGGAGG 5040
Qy 5041 TTCATACAGAAATTTATTTCAAGATGTACAGTTTCTACAAATGATGAAGAGCAGAGGCT 5100
Db 5041 TTCATACAGAAATTTATTTCAAGATGTACAGTTTCTACAAATGATGAAGAGCAGAGGCT 5100
Qy 5101 GTGTGGCAACACCAAGCAACCTATATGTTGATACATCTGCTGAATCTCGACCTCAT 5160
Db 5101 GTGTGGCAACACCAAGCAACCTATATGTTGATACATCTGCTGAATCTCGACCTCAT 5160
```

```
Qy 5161 ATTTATGATGATATAGATACATATACTATATATTTTATTTGTGGATTTCTCTAAAGCC 5220
Db 5161 ATTTATGATGATATAGATACATATACTATATATTTTATTTGTGGATTTCTCTAAAGCC 5220
Qy 5221 TCAGAAAATAGGACTGACTAGGCGAGCAAGAGAGAGTATCTTCTGTACACTGTTCCGCA 5280
Db 5221 TCAGAAAATAGGACTGACTAGGCGAGCAAGAGAGAGTATCTTCTGTACACTGTTCCGCA 5280
Qy 5281 GTTACTGGTATCATGAACAGTTGGAACCTGCTGCTTCTTCTTCTTCTTCTTCTTCTC 5340
Db 5281 GTTACTGGTATCATGAACAGTTGGAACCTGCTGCTTCTTCTTCTTCTTCTTCTTCTC 5340
Qy 5341 TCCTTTTGTAGCCCTTTTGTAGGGGTTTCATGATTCATTTACCAAGTTTTCAGT 5400
Db 5341 TCCTTTTGTAGCCCTTTTGTAGGGGTTTCATGATTCATTTACCAAGTTTTCAGT 5400
Qy 5401 TACCATTGTATGCAAGAGCCAGCAGTGAATACCTACATAGGTTTCTTCTTCTTCTTCT 5460
Db 5401 TACCATTGTATGCAAGAGCCAGCAGTGAATACCTACATAGGTTTCTTCTTCTTCTTCT 5460
Qy 5461 TTTTAAAGAGCGTAAATGACAGTGGAAACATAATGGGATATGAGAGACACCTTCAAGTT 5520
Db 5461 TTTTAAAGAGCGTAAATGACAGTGGAAACATAATGGGATATGAGAGACACCTTCAAGTT 5520
Qy 5521 ATTTCTGAATGATTTTGTAGGGTAAATATACAGATGCTTCTTGTATTAACAACTTGTGGA 5580
Db 5521 ATTTCTGAATGATTTTGTAGGGTAAATATACAGATGCTTCTTGTATTAACAACTTGTGGA 5580
Qy 5581 AAGCAGGAATCAGTGTCTTAAAGCTGTCATCTTATTAACCAATGCGGTTGTGCTATAAC 5640
Db 5581 AAGCAGGAATCAGTGTCTTAAAGCTGTCATCTTATTAACCAATGCGGTTGTGCTATAAC 5640
Qy 5641 TGCTGCTGATTTAGAGGGGAAC 5662
Db 5641 TGCTGCTGATTTAGAGGGGAAC 5662
```

```
RESULT 3
LOCUS CO869396 5127 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 2 from Patent WO2004074448.
ACCESSION CO869396
VERSION CO869396.1 GI:51999315
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Sabatini, D.M., Kim, D.H. and Sarbassov, D.D.
TITLE MTOR KINASE-ASSOCIATED PROTEINS
JOURNAL Patent: WO 2004074448-A 2 02-SEP-2004;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES
source
1. 5127
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

```
Query Match 90.6%; Score 5127; DB 6; Length 5127;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 28 ATGGCGCGCATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC 87
Db 28 ATGGCGCGCATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC 60
Qy 88 AGCGGCGAGGAGAACCTCCCGCTGGATCTGACCCGAGAACCTTCTGATACTTAAGAG 147
Db 61 AGCGGCGAGGAGAACCTCCCGCTGGATCTGACCCGAGAACCTTCTGATACTTAAGAG 120
Qy 148 ATTCTCCAAAATGTGCCAGATTGTCAGGGAGTATCAAAATATGAGAAAGCTTAGGCCATCTG 207
```

Db ||||| 121 ATTCTCCAAAATGTGGCCAGATTGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG 180
QY ||||| 208 AATAACTTTACTAAGCTTCTTTGTGATATTGGCCACAGTGAAGAAAACCTGGCTTTTAC 267
Db ||||| 181 AATAAATTTACTAAGCTTCTTTGTGATATTGGCCACAGTGAAGAAAACCTGGCTTTTAC 240
QY ||||| 268 TATGAGATATCATAAATTTGTTGGGTTAGCTTTTATTAATGAAGCAAAAGTGGCA 327
Db ||||| 241 TATGAGATATCATAAATTTGTTGGGTTAGCTTTTATTAATGAAGCAAAAGTGGCA 300
QY ||||| 328 GCAGCAGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGTATTCCTCAGAAAGGTG 387
Db ||||| 301 GCAGCAGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGTATTCCTCAGAAAGGTG 360
QY ||||| 388 CTAAATTTGAAGTGGACTATTTAATAGCTAGTGGTGCATGACATCAACAGAGCAACGAG 447
Db ||||| 361 CTAAATTTGAAGTGGACTATTTAATAGCTAGTGGTGCATGACATCAACAGAGCAACGAG 420
QY ||||| 448 GTAGAGAGACACAAGCACTTCGATTTAGTTCAGAAAGATGATTACTGTGAATCTTCCCTTG 507
Db ||||| 421 GTAGAGAGACACAAGCACTTCGATTTAGTTCAGAAAGATGATTACTGTGAATCTTCCCTTG 480
QY ||||| 508 TTTCTAGTTCGTGACCAACTCATTAATTTGAGTTGGAAATGATGGACTTCAAGAAAGA 567
Db ||||| 481 TTTCTAGTTCGTGACCAACTCATTAATTTGAGTTGGAAATGATGGACTTCAAGAAAGA 540
QY ||||| 568 GACAGATGGTCCGAGCATGCTATGCAATTTCTGTGAACACTGACCTTCAGAAATCCAGAG 627
Db ||||| 541 GACAGATGGTCCGAGCATGCTATGCAATTTCTGTGAACACTGACCTTCAGAAATCCAGAG 600
QY ||||| 628 GTGGTGGCCCTTCGAGGAGGACTTAAACCATATTGAAAAATGTGATTGATGCCAAATTA 687
Db ||||| 601 GTGGTGGCCCTTCGAGGAGGACTTAAACCATATTGAAAAATGTGATTGATGCCAAATTA 660
QY ||||| 688 AGTCGAATAAATGAGGCCCTTAATTACTACAAATTTTGACCTTTCTTAATCATCCAAAGACT 747
Db ||||| 661 AGTCGAATAAATGAGGCCCTTAATTACTACAAATTTTGACCTTTCTTAATCATCCAAAGACT 720
QY ||||| 748 CGGCAGTATGCGAGCTGATAGNAATTTAGAGAGNAATTTAGACCCCTATGATTTT 807
Db ||||| 721 CGGCAGTATGCGAGCTGATAGNAATTTAGAGAGNAATTTAGACCCCTATGATTTT 780
QY ||||| 808 CACTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAAAGAGACAGAGAAACGCA 867
Db ||||| 781 CACTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAAAGAGACAGAGAAACGCA 840
QY ||||| 868 TTTCTAGCCAGTAAATGGGAATCATAGCAACATTTCCGATCATGGGCGAGGTATTATTAAT 927
Db ||||| 841 TTTCTAGCCAGTAAATGGGAATCATAGCAACATTTCCGATCATGGGCGAGGTATTATTAAT 900
QY ||||| 928 TTATGTAAACCTGGAAATTTCTGGATCCAGTCTCTAATAGGAGTACTTTGCAATCAAAAT 987
Db ||||| 901 TTATGTAAACCTGGAAATTTCTGGATCCAGTCTCTAATAGGAGTACTTTGCAATCAAAAT 960
QY ||||| 988 ATGGAAATAAGCGAGGTCTACTTGAAGTCTTTATGATATATTCGCTCTCTCTACCT 1047
Db ||||| 961 ATGGAAATAAGCGAGGTCTACTTGAAGTCTTTATGATATATTCGCTCTCTCTACCT 1020
QY ||||| 1048 GTTGTGACTGAGGAGTTCATAGAAGCACTACTCAGTGTAGATCCAGGGAGGTTCCAAAGAC 1107
Db ||||| 1021 GTTGTGACTGAGGAGTTCATAGAAGCACTACTCAGTGTAGATCCAGGGAGGTTCCAAAGAC 1080
QY ||||| 1108 AGTTGGAGCTTCAGATGGCTTTGGCGAGCTGAGGCAAAACTATTTCTTCTCATCGT 1167
Db ||||| 1081 AGTTGGAGCTTCAGATGGCTTTGGCGAGCTGAGGCAAAACTATTTCTTCTCATCGT 1140
QY ||||| 1168 GCCAGATCCAGGCGCAGACCTCATGGATAATTTTGGCACTGATCTCTGCAATTTAT 1227
Db ||||| 1141 GCCAGATCCAGGCGCAGACCTCATGGATAATTTTGGCACTGATCTCTGCAATTTAT 1200
QY ||||| 1228 CGTAAATGGACTTTTATGAGGCTCTAGTTGAAGTGATAACAAACAGTGAATCATATCTCA 1287

Db ||||| 1201 CCTAATGCACTTTTAGAGGGTCTAGTTGAAGTGATAACAAACAGTGAATCATATCTCA 1260
QY ||||| 1288 GTTAGAGCTTACCATCTTTTAGGAGAGCTTTTACATATGGCAAAACACAAATTTCTTCTCAT 1347
Db ||||| 1261 GTTAGAGCTTACCATCTTTTAGGAGAGCTTTTACATATGGCAAAACACAAATTTCTTCTCAT 1320
QY ||||| 1348 TCACATAGCCATCATTTTACACTGCTTGCACCCCTTAATGAATATGGCTGCATCTTTTGAAT 1407
Db ||||| 1321 TCACATAGCCATCATTTTACACTGCTTGCACCCCTTAATGAATATGGCTGCATCTTTTGAAT 1380
QY ||||| 1408 ATCCCAAGGAAAAGAGAGCTGCGAGCCAGTGAGCCCTTGAACCTGTTTAAAGCTTCCAT 1467
Db ||||| 1381 ATCCCAAGGAAAAGAGAGCTGCGAGCCAGTGAGCCCTTGAACCTGTTTAAAGCTTCCAT 1440
QY ||||| 1468 GAAATGAAGAACGAGGACCTTAAGCCCTTATAGTCTTCAATTTAGACCAATATTTCAGAAA 1527
Db ||||| 1441 GAAATGAAGAACGAGGACCTTAAGCCCTTATAGTCTTCAATTTAGACCAATATTTCAGAAA 1500
QY ||||| 1528 GCAATTTGCAACACACACAGAAACGGGATCAGTATCTCCGAGTTTCAGAAAGATATATTTATC 1587
Db ||||| 1501 GCAATTTGCAACACACACAGAAACGGGATCAGTATCTCCGAGTTTCAGAAAGATATATTTATC 1560
QY ||||| 1588 CTTAAGGATACAGAGAAAGCTCTTTTAAATTAACCTTAGAGATAGCAAGCTCTTCAACAT 1647
Db ||||| 1561 CTTAAGGATACAGAGAAAGCTCTTTTAAATTAACCTTAGAGATAGCAAGCTCTTCAACAT 1620
QY ||||| 1648 AAAGAGAAATCTTGAATGGAATCTTATAGGAGCCATTTTAAAGTGGCCAAATGTA 1707
Db ||||| 1621 AAAGAGAAATCTTGAATGGAATCTTATAGGAGCCATTTTAAAGTGGCCAAATGTA 1680
QY ||||| 1708 AATCTTAAGAAACTATAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT 1767
Db ||||| 1681 AATCTTAAGAAACTATAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT 1740
QY ||||| 1768 TACAAGCCGACAGTAAATATATATGCCAACTGGAATCTGGAATTTTGCAGGCCCAACAG 1827
Db ||||| 1741 TACAAGCCGACAGTAAATATATATGCCAACTGGAATCTGGAATTTTGCAGGCCCAACAG 1800
QY ||||| 1828 CTCACGGTTGTAGGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGATGGGCA 1887
Db ||||| 1801 CTCACGGTTGTAGGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGATGGGCA 1860
QY ||||| 1888 GGCTACTTAGAAGATCTAGTAAAGGATATTTGTCAGTGGCTCAATGCTTCTTCTGGAATG 1947
Db ||||| 1861 GGCTACTTAGAAGATCTAGTAAAGGATATTTGTCAGTGGCTCAATGCTTCTTCTGGAATG 1920
QY ||||| 1948 AAACCCGAAAGAGTCTTCAAAATTAATGGTTTATGACCACTTGTAGTCAACACTACTTT 2007
Db ||||| 1921 AAACCCGAAAGAGTCTTCAAAATTAATGGTTTATTTGACCCCTTTAGTCAACACTACTTT 1980
QY ||||| 2008 TTATTTTATTTGGAACACTTTTCCACCCCTCATGAGTTTAAATGCTGGAATAATGCAATG 2067
Db ||||| 1981 TTATTTATTTGGAACACTTTTCCACCCCTCATGAGTTTAAATGCTGGAATAATGCAATG 2040
QY ||||| 2068 GTATTTTCACTGCTCTTCTTAAATCTTTGCTCTTTGAAAAACCAAGATCACTTGTCAAAACT 2127
Db ||||| 2041 GTATTTTCACTGCTCTTCTTAAATCTTTGCTCTTTGAAAAACCAAGATCACTTGTCAAAACT 2100
QY ||||| 2128 ACTGTTTCTAGCTTGGACTATAGCAGAGATGGATGGCTTAGGTCATCTCTTCCAAAT 2187
Db ||||| 2101 ACTGTTTCTAGCTTGGACTATAGCAGAGATGGATGGCTTAGGTCATCTCTTCCAAAT 2160
QY ||||| 2188 TTAACCTGAGCTTACTGATGCTTCCAGACTCTATGCAACAAACATTTAAGGGTATTATG 2247
Db ||||| 2161 TTAACCTGAGCTTACTGATGCTTCCAGACTCTATGCAACAAACATTTAAGGGTATTATG 2220
QY ||||| 2248 AGAGCTAATTTGAAATCTTTTAAATTTGGGAAATTTAGTTTGTAGTACCAGCTACAT 2307
Db ||||| 2221 AGAGCTAATTTGAAATCTTTTAAATTTGGGAAATTTAGTTTGTAGTACCAGCTACAT 2280
QY ||||| 2308 GATAAAAAACAAACGAAATTTCTTGAAGCTCTTGAATCTTCTCGATGAAGCATGTGAAGAC 2367
Db ||||| 2281 GATAAAAAACAAACGAAATTTCTTGAAGCTCTTGAATCTTCTCGATGAAGCATGTGAAGAC 2340

QY 2368 AAGGCCAATCTTCATGCTCTCATTCAGATGAAACGAGGTTATCCCACTTGGAGACAAG 2427
DB AAGGCCAATCTTCATGCTCTCATTCAGATGAAACGAGGTTATCCCACTTGGAGACAAG 2400
QY 2428 GGTTCCTTCCTCCGCTGAGATTTCTCTCCANTCCAAAGGATTTCTCTATCTGAATGAA 2487
DB GGTTCCTTCCTCCGCTGAGATTTCTCTCCANTCCAAAGGATTTCTCTATCTGAATGAA 2460
QY 2488 AGAGTTATGTAGCAAAACAAATTTGAAAAGTGGCACAGGGAATACAACTCCAAATATGTT 2547
DB AGAGTTATGTAGCAAAACAAATTTGAAAAGTGGCACAGGGAATACAACTCCAAATATGTT 2520
QY 2548 GACTTGTATGAGAAACAACTCAATGAAGCACTTACTACTTACCCGGAAGCCCTGTTGATGGT 2607
DB GACTTGTATGAGAAACAACTCAATGAAGCACTTACTACTTACCCGGAAGCCCTGTTGATGGT 2580
QY 2608 GATAACTATGTTCTCGGAGTAACCAAGATTTACAGGCTCTCAAGCTTACCTGCCTATA 2667
DB GATAACTATGTTCTCGGAGTAACCAAGATTTACAGGCTCTCAAGCTTACCTGCCTATA 2640
QY 2668 CACCTTTATGACAACTAGTACACATATAAAACAGGCTGCCATTTGTTGGAAGTACAGAT 2727
DB CACCTTTATGACAACTAGTACACATATAAAACAGGCTGCCATTTGTTGGAAGTACAGAT 2700
QY 2728 ATTATTAAGAACTCTCTCGTAATGTTTCGTACACCAAGATTTGGATAAGTGGGAAGAAAT 2787
DB ATTATTAAGAACTCTCTCGTAATGTTTCGTACACCAAGATTTGGATAAGTGGGAAGAAAT 2760
QY 2788 AAAAACTGAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAATTTGGGTC 2847
DB AAAAACTGAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAATTTGGGTC 2820
QY 2848 AATTGCTACAGGAAGAAACGTAATTCAGATATCTAAAACTTGCAAAACAGTGTGAA 2907
DB AATTGCTACAGGAAGAAACGTAATTCAGATATCTAAAACTTGCAAAACAGTGTGAA 2880
QY 2908 GTTCTTTCCATCAGAGGAACTGTGTATATGTACTTGGGCTCATAGCTAAAAACCAACAA 2967
DB GTTCTTTCCATCAGAGGAACTGTGTATATGTACTTGGGCTCATAGCTAAAAACCAACAA 2940
QY 2968 GGCCTGTATATCTTAAATGTCACACTGGGATGCTGTGGGATAGTCGCAACATCTG 3027
DB GGCCTGTATATCTTAAATGTCACACTGGGATGCTGTGGGATAGTCGCAACATCTG 3000
QY 3028 TGGCCAGTGTTCAGATGATGTGGAACAACTCTGTAATGAACCTTCATCTATCCCAAGC 3087
DB TGGCCAGTGTTCAGATGATGTGGAACAACTCTGTAATGAACCTTCATCTATCCCAAGC 3060
QY 3088 ACTTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATTAATAGTGAAGTGAATCTGTG 3147
DB ACTTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATTAATAGTGAAGTGAATCTGTG 3120
QY 3148 CCATCGAGTATGTATATTTGGAGGATGACGGTTTGGAGGAGCTCTACTAGTACATTT 3207
DB CCATCGAGTATGTATATTTGGAGGATGACGGTTTGGAGGAGCTCTACTAGTACATTT 3180
QY 3208 TTCTCTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGACCCCAATAAG 3267
DB TTCTCTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGACCCCAATAAG 3240
QY 3268 GATAAAATTCATTCCTTTCTTTGCTTCTAGTAAACCTTGTGAAGATCGTATCTTAAAT 3327
DB GATAAAATTCATTCCTTTCTTTGCTTCTAGTAAACCTTGTGAAGATCGTATCTTAAAT 3300
QY 3328 TCGTTACTTTGGCTAAACAAAACATCGTAGTAGAGTATCCAAAGAGGGAATTA 3387
DB TCGTTACTTTGGCTAAACAAAACATCGTAGTAGAGTATCCAAAGAGGGAATTA 3360
QY 3388 TCATCTGAAGTAAAGCAACGAGCGAATCAGAACCTTACGGAGCCAGTGTGAT 3447
DB TCATCTGAAGTAAAGCAACGAGCGAATCAGAACCTTACGGAGCCAGTGTGAT 3420

QY 3448 TTTAATCATAGTATGATATTTTACACCCATATCCACTGTACAGAAAACATTTAATAGAG 3507
DB TTTAATCATAGTATGATATTTTACACCCATATCCACTGTACAGAAAACATTTAATAGAG 3480
QY 3508 ACTTCATTTATGGGATAGCACATTTGAAGACACTGGTAGTACACCAAGCATTTGGAGAA 3567
DB ACTTCATTTATGGGATAGCACATTTGAAGACACTGGTAGTACACCAAGCATTTGGAGAA 3540
QY 3568 AATGACTTTAAATTTACCAAGAAATTTTGGTACAGAGAAATACAGAGAAATACAGCCGA 3627
DB AATGACTTTAAATTTACCAAGAAATTTTGGTACAGAGAAATACAGAGAAATACAGCCGA 3600
QY 3628 GAGAGTTAGTATGAGAAAGTTTCAACGAGTCTACATATGAAGATACGTAGCCAAAGTTTC 3687
DB GAGAGTTAGTATGAGAAAGTTTCAACGAGTCTACATATGAAGATACGTAGCCAAAGTTTC 3660
QY 3688 AATACAGACACTACAAAGTGGGATAGTTCAATGAGCTCAAGTCTTCCAGAGAGACA 3747
DB AATACAGACACTACAAAGTGGGATAGTTCAATGAGCTCAAGTCTTCCAGAGAGACA 3720
QY 3748 GTAGGTAGTATGCTACAACTATGGGACACAGACTGTGGAAAGCATGAGTACTGTGTAAGT 3807
DB GTAGGTAGTATGCTACAACTATGGGACACAGACTGTGGAAAGCATGAGTACTGTGTAAGT 3780
QY 3808 ACTAAAACTAATTAAGACAAAGCCACTATTTGACGCGCAAGTCTAACCATCTCTCTCC 3867
DB ACTAAAACTAATTAAGACAAAGCCACTATTTGACGCGCAAGTCTAACCATCTCTCTCC 3840
QY 3868 AATCAAAATTCGGTGTCCCTGGTCCCTCAGAGTTCTTCTCATAGCTTCTTAGAAGACA 3927
DB AATCAAAATTCGGTGTCCCTGGTCCCTCAGAGTTCTTCTCATAGCTTCTTAGAAGACA 3900
QY 3928 CAGTCCCTTAAAGCACCCTCTATTGTCTACAAATTTAAAGTCTAGCAGATTTGTAATTTAGT 3987
DB CAGTCCCTTAAAGCACCCTCTATTGTCTACAAATTTAAAGTCTAGCAGATTTGTAATTTAGT 3960
QY 3988 TACAAGTTCTAGAGATGCTTTTGGCTATGTCTACACTGAAAAAGACTACAGCAACAAGA 4047
DB TACAAGTTCTAGAGATGCTTTTGGCTATGTCTACACTGAAAAAGACTACAGCAACAAGA 4020
QY 4048 ATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4107
DB ATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4080
QY 4108 ACTGATACCATCAACATGAAGGCCAACAGTTTGTAGTCCAGATTTAAACCAAGCAGTTTC 4167
DB ACTGATACCATCAACATGAAGGCCAACAGTTTGTAGTCCAGATTTAAACCAAGCAGTTTC 4140
QY 4168 ATGAAAGCCTTAAGTTATGCATCATTTAGATATAAGAAAGATTTATTTAGTCTCTTAAATCAA 4227
DB ATGAAAGCCTTAAGTTATGCATCATTTAGATATAAGAAAGATTTATTTAGTCTCTTAAATCAA 4200
QY 4228 AATACCTGCAAGCATCTTCTCAGTGGGTCCATGGTGTCCAGTGGCCACATATGGGGGT 4287
DB AATACCTGCAAGCATCTTCTCAGTGGGTCCATGGTGTCCAGTGGCCACATATGGGGGT 4260
QY 4288 TCAGATGATTTACATTTGCTCTCTCCGCTGATATAAATGATATATTTCCAGGTAAAG 4347
DB TCAGATGATTTACATTTGCTCTCTCCGCTGATATAAATGATATATTTCCAGGTAAAG 4320
QY 4348 GATATTTCCCTTTTTCAGACAAAACATACCAACCATGATGATCGAGGTGCAAGAGCA 4407
DB GATATTTCCCTTTTTCAGACAAAACATACCAACCATGATGATCGAGGTGCAAGAGCA 4380
QY 4408 TTTGCCCATGATGCGAGGCTCTTCATCTGGAACTGGAGGTCTTTGTAATAAATCTCTTTT 4467
DB TTTGCCCATGATGCGAGGCTCTTCATCTGGAACTGGAGGTCTTTGTAATAAATCTCTTTT 4440
QY 4468 CACTTCTCAGCAGCAGATGAGTCTTACGGAAATTAATGAATTCATTCATTCAGATGCC 4527
DB CACTTCTCAGCAGCAGATGAGTCTTACGGAAATTAATGAATTCATTCATTCAGATGCC 4500
QY 4528 TCTCTGTTTTTAGAAAAGTACAGAAAGACACTGGGACTACAGGAACATACAGATGATACTGC 4587

Db	4501	TCCTGTTTTAGAAAGTACAGAAGACACTGGACTACAGAACATACAGATGAATACTGC	4560
Qy	4588	CTTTATTGTGTCGTATTGAAATTTCTGGGTTTCCAGCCGACCAACCACTGAGTGC AATA	4647
Db	4561	CTTTATTGTGTCGTATTGAAATTTCTGGGTTTCCAGCCGACCAACCACTGAGTGC AATA	4620
Qy	4648	TGTAGTCATTCCAGACCTTTCAAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCAT	4707
Db	4621	TGTAGTCATTCCAGACCTTTCAAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCAT	4680
Qy	4708	AATCCCTTTAGAAGTGGTTCCTCTAAAGTTTTCGGGGATTTCTGGATCGACGTATGGGGTG	4767
Db	4681	AATCCCTTTAGAAGTGGTTCCTCTAAAGTTTTCGGGGATTTCTGGATCGACGTATGGGGTG	4740
Qy	4768	TCTCAAGAAGGCTCAGCTAGCAGCAGCA CAAAGAACA GAGAATTTGTTACTAGGTTGTTAAACA	4827
Db	4741	TCTCAAGAAGGCTCAGCTAGCAGCAGCAGCA CAAAGAACA GAGAATTTGTTACTAGGTTGTTAAACA	4800
Qy	4828	ATTCCAGATGATACACCAAATGTGCGGTATACTCTCTCGCAAGAAGTTCTCAAGATTAGTC	4887
Db	4801	ATTCCAGATGATACACCAAATGTGCGGTATACTCTCTCGCAAGAAGTTCTCAAGATTAGTC	4860
Qy	4888	ATTAAATTTGAGTAGTTTCAGTTTTCAACTAAATGATGAGACTGGGCTTTTAAACAATTAAG	4947
Db	4861	ATTAAATTTGAGTAGTTTCAGTTTTCAACTAAATGATGAGACTGGGCTTTTAAACAATTAAG	4920
Qy	4948	GAGAAGTATCCTCAAAACATTTGATGACATATGCTTTTACTCTGAGGTTTCCCAATTTGCTG	5007
Db	4921	GAGAAGTATCCTCAAAACATTTGATGACATATGCTTTTACTCTGAGGTTTCCCAATTTGCTG	4980
Qy	5008	TCACTATGCACTTCAGACTTCGGTTCGGAGTTTCATACAAGAAATTTATTTCAAGATGTA	5067
Db	4981	TCACTATGCACTTCAGACTTCGGTTCGGAGTTTCATACAAGAAATTTATTTCAAGATGTA	5040
Qy	5068	CAGTTTTCTACAAATGCATGAAGAAGCAGAGGCTGTGTTGCGACACACCAAGACCAACCT	5127
Db	5041	CAGTTTTCTACAAATGCATGAAGAAGCAGAGGCTGTGTTGCGACACACCAAGACCAACCT	5100
Qy	5128	ATAGTTGATACATCTGCTGAATCCTGA	5154
Db	5101	ATAGTTGATACATCTGCTGAATCCTGA	5127

```

RESULT 4
AY540053
LOCUS      9297 bp     mRNA     linear     ROD 03-MAR-2004
DEFINITION Mus musculus pianissimo mRNA, complete cds.
ACCESSION AY540053
VERSION   AY540053.1  GI:44241266
KEYWORDS  .
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 9297)
AUTHORS   Shiota,C., Woo,J.T., Lindner,J., Shelton,K.D., Zhu,M.Y. and
            Magnuson,M.A.
TITLE      Mouse pianissimo
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 9297)
AUTHORS   Shiota,C., Woo,J.T., Shelton,K.D. and Magnuson,M.
TITLE      Direct Submission
JOURNAL    Submitted (03-FEB-2004) Molecular Physiology and Biophysics,
            Vanderbilt University, 21st and Garland Avenues, Nashville, TN
            37232-0615, USA

FEATURES
            Location/Qualifiers
                1..9297
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="LAF1"
                    /db_xref="taxon:10090"
                    /chromosome="15"
source

```


Db 361 |||||CCAGTATTTCTCCAGAGGTGCTAAAACTGAAAGTGGGCTATTTTAATAGCTAGGTGCAATTG 420
Qy 428 ACATACAAACAGAGCAACGAGGTAGAGAGGACACAAGCACCTTCGATTAGTCAGAAAGATGA 487
Db 421 ACATTCAGCAGAGCAACGAGGTGAGAGGACACAAGCCCTTCGCTTAGTCAGAAAGATGA 480
Qy 488 TTACTGTGAATGCTTCCTGTTCTTCTAGTTCCTAGTTCCTGACCAACTCATTAATTCGAGTTGGAA 547
Db 481 TTACTGTGAATGCTTCCTGTTCTTCTAGTTCCTGAGCCAACTCATTAATTCGCGTTGGAA 540
Qy 548 ATGATGACCTTCAGAAAGAGACAGCAAGATGGTCCGAGCATGCAATTCGCTGAAC 607
Db 541 ATGATGGGCTTCAGAAACGAGACAGAAATGGTCCGCGGCTTCGATCCCATCATCTGTGAAC 600
Qy 608 TAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCAGAGGAGCACTAAACACCATATTTGAAAA 667
Db 601 TAGCACTTCAGAACCCAGAGGTGGTGGCCCTTCAGAGGAGCACTAAACACCATATCTGAAAA 660
Qy 668 ATGTGATGTGATGTCGAATTAAGTCGAATAAATGAGGCCCTTAATPACTACAAATTTGCAAC 727
Db 661 ATGTGATGTGATGTCGAATTAAGTCGAATAAATGAGGCTCTCATTTACTACAAATTTGCAAC 720
Qy 728 TTCTTAATCATCTCAAGACTCGGCAGTATGTGCGAGCTGATGTAGAATTTAGAGAAATTT 787
Db 721 TTCTTAATCATCTCAAGACCGGCAGTATGTGCGAGCTGATGTAGAATTTAGAGCGAATTT 780
Qy 788 TAGCACCTTACTACTGATTTTCACTPACAGACATAGTCCAGATACAGCTGAAAGGACAGCTCA 847
Db 781 TAGCACCTTACTACTGATTTTCACTPACAGACATAGTCCAGATACAGCTGAAAGGACAGCTCA 840
Qy 848 AAGAAGACAGAAAGACGATTTCTAGCCAGTAAATGGAATCATAGCAACATTCGGAT 907
Db 841 AAGAAGACAGAAAGACGATTTCTAGCCAGTAAATGGAATCATTTGCGACATTCGGCT 900
Qy 908 CATGGGAGGTATTAATTAATTAATGTAATCTGGAATTTCTGGATCCAGTCTCTAATAG 967
Db 901 CGTGGGAGGTATTAATTAATTAATGTAATCTGGAATTTCTGGATACAGTCTCTAATAG 960
Qy 968 GAGTACTTTGCATACCAATATGAAATAGAGGAGGCTACTTTGAAAGTCTTTATGATA 1027
Db 961 GAGTACTTTGCATACCAATATGAAATAGAGGAGGCTCTGCTGAAAGTCTCTATGACA 1020
Qy 1028 TATTTGCTCTCTCTACCTGTTGTGACTGAGGAGTTTCATAGAGCACTACTCAGTGTAG 1087
Db 1021 TCTTTGCTCTCTCTGCTGCTGAGCGATGAGTTTCATAGAGCCCTGCTCAGTGTAG 1080
Qy 1088 ATCCAGGAGGTTCCAAAGACAGTTGGAGGCTTTTCAGATGGCTTTGGCGAGCTGAGGCAA 1147
Db 1081 ATCCAGGAGGTTCCAGGACAGTTGGAGGCTTTCCGATGGCTTTGTTGGCAGCTGAGGCAA 1140
Qy 1148 AACTATTTCTCTCATCGTCCAGATCCAGGCCAGACCTCATGGATAATTTATGGCAC 1207
Db 1141 AACTATTTCTCTCATCGTCCAGTAGATCCAGGCCAGACCTCATGGCAATTTATGGCTC 1200
Qy 1208 TGATACTCTCTGCAATTTATTCGTAATGCACTTTTATAGAGGCTCTAGTTGAAAGTATATAAA 1267
Db 1201 TGATACTCTCTGCAATTTATTCGTAATGCACTTTTATAGAGGCTTGTCTGAGGTGATAACGA 1260
Qy 1268 ACAGTATGATCATATCTCAGTTAGAGCTACCATTCCTTTTATAGGAGCTTTTACATATGG 1327
Db 1261 ACAGTATGAGCCATCTCAGTATGAGGCGACCATTCCTTTTATAGGAGCTCTCTGCATATGG 1320
Qy 1328 CAACACAAATTTCTCTCATCTCATAGAGCCATCATTTTACACTGCTGGCAACCTTATGA 1387
Db 1321 CAACACAAATCTCTGCTCATCTGATAGCCACCATCTACACTGCTTGGCAACCTTATGA 1380
Qy 1388 ATATGGCTGATCTTTGATATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCCCTTGA 1447
Db 1381 ACATGGCTGCTCTTCGATATTTCCCAAGGAAAGAGGCTTGGAGCCAGTGGAGCCCTGA 1440
Qy 1448 ACTGTTTTAAACGCTTCCATGAATTAAGAAACGAGGACCTTAAGCCCTTATAGTCTTCATT 1507
Db |||||

1441 ACTGTTTGAACCCGCTTCCATGAGATGAAGAAGCGAGGACCTAAGCCCTCAGCGCTTCATT 1500
1508 TAGACCACTATTTTCAGAAAGCAATTTGCAACACACACAGAAACGGGATCAGTATCTCCGAG 1567
1501 TAGATCACATCATTTAGAAAGCAATCGCAACTCACCACAAGCGGGATCAGTATCTTCGAG 1560
1568 TTCAAGAAAGATATATTTATCTTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAG 1627
1561 TTCAAGAAAGATATATTTGTTCTTTAAGGATACAGAGGAGCTCTTTTAAATAAACCTTAGAG 1620
1628 ATAGCCAAAGTCTTCACATTAAGAGAACTCTTAAGTGAATTCGAAATCTGATTTGGGACCA 1687
1621 ACAGCCAAAGTCTTCAGCATTAAGAGAACTCTTCACTGGGATTTGAAATCTGATTTGGGACCA 1680
1688 TTCTTAAGTGGCCAAATGTAATCTAAGAACTATATAAGATGAACAGTATACACAGGTTTG 1747
1681 TCCTTAAGTGGCCAAATGTAATCTAAGAACTATATAAGATGAGCAGTATGACACAGGTTTG 1740
1748 TAGAAGACTACTTTTATTTTACAAGCCAGCAGTAGTAAATTAATATATGCAACCTTGATCTGG 1807
1741 TGGCAGACTCTCTTACTTTTACAAGCCAGCAGCAAACTGTACGCTAGTCTGGATCTGG 1800
1808 ATTTTGGCAGGCAACAGCTCAAGGTTGTAGGTTGCCAGTTCACAGAAATTTCTTCTTG 1867
1801 ACTTGGCCAGTCCAAGCAGCTCACAGTTGTGCGTTGTAGTTTACAGAAATTTCTGCTCG 1860
1868 AATCTGAAGAGGTGGCAAGGCTACTTGAAGATCTAGTAAAGGATATTTGTTCAAGTGGC 1927
1861 AGTCTGAAGAGGTGGCAAGGATCTTAGAAGATCTCGTGAAGATATTTGTTCAAGTGGC 1920
1928 TCAATGCTTCACTCGGAATGAAACCCGAAAGAGTCTTCAAAAATAATGTTTATTTAGCA 1987
1921 TCAATGCTTCACTCGGAATGAAACCTTGAGAGAGCTTTCAGAAATATGTTGCTTACTGACCA 1980
1988 CCCTTAGTCAACTACTTTTATTTATTTGGAACACTTTCTTGCACCCCTCATGGAGTTA 2047
1981 CCCTTAGTCAACTACTTTTATTTATTTATCGGAACTCTCTTGTCTCATCTCATGGAGTCA 2040
2048 AATGCTGGAAGAAATGCAAGTATTTTCAAGTCTCTTAAATCTTGTCTCTCGTGAAGAAC 2107
2041 AATGCTGGAAGAAATGCAAGTCTCTTAAATCTTGTCTCTTAAATCTTGTCTCTGGAAGAAC 2100
2108 AAGATCACTTGTCTTAAACTTACTGTTTCTAGCTTTGAGCTATATAGCAGAGATGGATGGCTA 2167
2101 AAGATCACTTGTCTTAAACTTACTGTTTCTAGCTTTGAGCTATATAGCAGAGATGGATGGCTA 2160
2168 GAGTCACTTCTTCAAAATTTTAACTGCAAGTACTGATGCTGCAAGTCTCTATGCAACAA 2227
2161 GAGTCACTTCTTCAAAATTTTAACTGCAAGTACTGATGCTGCAAGTCTCTATGCAACAA 2220
2228 AACATTTAAGGTTATTTATGAGAGCTAAATGTTGAAATTTCTTTAATAATTTGGGAAATTCAGT 2287
2221 AACATTTAAGGTTATTTATGAGAGCTAAATGTTGAAATTTCTTTAATAATTTGGGAAATTCAGT 2280
2288 TGTAGTGAACCCAGCTACATGATATAAACAACAAATTCCTCTGAGGCTCTTGTATATCC 2347
2281 TACTAGTAACTCAGCTACATGATATAAACAACAAATTTCTTCTGAGCGCTGATATTC 2340
2348 TCATGAGAGCATGTGAAGCAAGGCCAAATCTTCAATGCTCTCATTCAGATGAAGAACCGCT 2407
2341 TTGATGAAGCTTGTGAAGCAAGGCCAAATCTTCAATGCTCTCATTCAGATGAAGAACCGCT 2400
2408 TATCCCACTTGGAGCAAGGGTTTCTCTCTGCTGAGATTTCTCTCCATTTCCAAAAG 2467
2401 TATCCCACTTGGAGCAAGGGTTTCTCTCTGCTGAGATTTCTCTCCATTTCCAAAAG 2460
2468 GATTTTCTCTATCTGAATGAAGAGGTTATGTAGCAAAAACAAATTTGAAAAGTGGCACAGGG 2527
2461 GATTTTCTCTATCTGAATGAAGAGGTTATGTGCAAAAACAGTTTGGAAAAGTGGCACAGGG 2520
2528 AATACAACTCCAAATATGTTGACTTGAATGAGGAAACAACTCAATGAAGCACTTACTACTT 2587
2521 AATATAATTCAAAGTATGTTGACTTGAATGAGGAAACAACTTAATGAAGCCCTCACTACAT 2580


```
Db 4738 GTGGCTGAGGATGGCCATCTCAGGAAGAGGCTCAGCAGCAGCAGCAAAAAGCACTG 4797
Qy 4805 AATTGTTACTAGGTGTTAAACAATTCAGATGATACACCAATGTCGCGTATCTCTCTTC 4864
Db 4798 AGCTGTTACTGGTGTAAACACATTCAGATGATACACCAATGTCGCGATCTGCTTC 4857
Qy 4865 GCAAAGAGTCTTAAGATTAGTCATTAATTTGAGTGTAGTTCAGTTCAACTAAATGTCTATG 4924
Db 4858 GTAAGGAAGTCTAAGATTAGTGTCAATTTGAGTGTAGTTCAGTTCACCAAGTGTCAAG 4917
Qy 4925 AGACTGGCTTTTACCAATTAAGGAGAGTATCTCAACATTTTGTATGATGACATATGCTTTT 4984
Db 4918 AGACGGGCTTCTGACCAATTAAGGAAGATGATCTCAACATTTTGTATGACATATGCTTTT 4977
Qy 4985 ACTCTGAGTTCCTCATTTGCTGTCACACTGCACATTCAGACTTCGCGTGTGCGAGGTTC 5044
Db 4978 ATTCTGAAGTTTCTCATCTGTTGTCACACTGCACATTTAGACTTCAGTGTGCGGCTTC 5037
Qy 5045 TACAAGATTAATTTCAAGATGTACAGTTCCTCAAAATGTCATGAAGAGCAGAGCTGTGT 5104
Db 5038 TACAAGAGTATTTTCAAGATGTACAGTTCCTCAAAATGTCATGAAGAGCAGAGCTGTGT 5097
Qy 5105 TGGCAACACCAACCAACCTATAGTGTATGATATCATCTGCTGAATCTCTGACCTCATATTT 5164
Db 5098 TGGCAATTTCCCAACCAATTCAGCTATAGTGTATGATGATGCTGCTGAATCTCTGACCTCATGCTT 5157
Qy 5165 ATGATGATATAGATACATATATATATATATATATATATATATATATATATATATATATAT 5223
Db 5158 ATGACGTTTATAGCTGATATCTATATATATATATATATATATATATATATATATATATAT 5217
Qy 5224 GAAATATACGACTGAGCAGCAAGCAGAGAGTATCTCTGATACACTGTTCGCGAGTT 5283
Db 5218 CAGGATATGACTGCTGACAG-TGAGCAGAGAGAGCTTCTG-ACACTGCGCAGAGTT 5275
Qy 5284 ACTGATACGACAGTGTGAGTCTGCTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5343
Db 5276 ATTGGTACA-GAACAGCTGGGAGCTGCTGGCAATCTCTAGCAGCAAGCAA-CTTCTCTCTC 5330
Qy 5344 TTTCTGAGCTTTTGTAGGGTTCATGATTCATTAACACAGTTCATTAAGATTTTCAGTTAC 5403
Db 5331 CGGTTGGCCCTTCTCGGGTTCGTTGTTTATTAACATGTTTAA-CTTCTCTCTC 5386
Qy 5404 CATTTGATGCAAGAGCAGCACTGAATACCTATAGTGTTCATTTCTTCTTCTTCTTCTTCTTCT 5463
Db 5387 CATCATATGCAAGAGCAGCACTGAATATCTATAGTGTTCATTT-CTTCTCTCTC 5444
Qy 5464 AAAAGCGTATGACAGTGTGAACAATATGAGATGATGAGAGCACTTCAACA-AGTTA 5521
Db 5445 CAAAGCGCAATGGCAGTGGAAACA-AAACAGGACATGAAGAGCATCTTCAACACCGTGT 5503
Qy 5522 TTTCTGATGATTTT-----AGGTAATAATACAGATGCTGTATGTTAACTACTT 5575
Db 5504 TTTCTGATGCTTTTAAATCAAGGGTAAATACAAAGATGCTGTGTTGTTAACTAAATTT 5563
Qy 5576 GTGAAAGCAGGAATCAGTGTCTTAAGGCTGATCTTATTAACCAATGG 5626
Db 5564 GTGAAAGCAGGAGTCACTGACTCGAGACTGATCTGTTACCCCGTGGG 5614
```

```
RESULT 5
AB082530
LOCUS AB082530 Homo sapiens mRNA for KIAA1999 protein. linear PRI 04-JUL-2002
DEFINITION Homo sapiens mRNA for KIAA1999 protein.
ACCESSION AB082530
VERSION AB082530.1 GI:21693143
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ohara,O., Nagase,T., Mitsui,G., Kohga,H., Kikuno,R., Hiraoka,S.,
```

Takahashi,Y., Kitajima,S., Saga,Y. and Koseki,H.
Characterization of size-fractionated cDNA libraries generated by
the in vitro recombination-assisted method
DNA Res. 9 (2), 47-57 (2002)
22051173
12056414
2 (bases 1 to 8213)
Direct Submission
Ohara,O., Nagase,T. and Kikuno,R.
Submitted (26-MAR-2002) Oeamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kibara, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)

FEATURES

source

1..8213
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="bf00346"
/tissue_type="brain"
/notes="vector:pBluescriptII SK plus"

gene

1..8213
/gene="KIAA1999"<1..3829
/gene="KIAA1999"/notes="Start codon is not identified. bf00346 cDNA clone
for KIAA1999 has a 19-bp insertion after the position 1492
of the sequence of KIAA1999."/codon_start=2
/evidence=not experimental
/product="KIAA1999 protein"

/protein_id="BAC02708.1"

/db_xref="GI:21693144"

/translation="ANTLPHSHSHLHCLPTLNNMAASPDIPKPKRLASALNCLK
RFLMKRGKPKYSLHLDHIQKAIATKQKQDQYLRQKDFIKDTKEALLNLRDS
QVQHEKLEWNNLITGLIKWPNVNDQYLRQKDFIKDTKEALLNLRDS
DPAKQYLVGQCTFFELLESDEGQVLDKDIQVNLNASSGKPKRSKLNQGL
LTTLSOHYFLFICTLSCHPHGVKLEKSVFQCLNLCSLKNODHLLKLTSSIDYSL
DGLARVLSKILTAADACRLYATKHLRVLLRANVEFPNNGWIELLVQLDKNKTIS
SEALDILDEACEDKANHLAIQKPAHSLGDKGLLLRLFLSPKGSYLNBERGYVA
KQLEKWHREYNSKYVDLIQKPAHSLGDKGLLLRLFLSPKGSYLNBERGYVA
QVQHEKLEWNNLITGLIKWPNVNDQYLRQKDFIKDTKEALLNLRDS
LQLEENVIQDILKQCEVLSIRGTCVVLGIATKQKQDILKCHNWDVARSKH
LWPPVDVQVQCNELSSIPSTLSANSESTSSRHNSSESVPSPMTLEDRGSSST
STFLDINDEPTFFYDRGSPIDKXNSPFPASSKLVRILNLSLILPKKRRSSSDP
GKGLSSSEKTSNRRILTEPSVDFNHSDDFTPISTVQKTLQETSFMGNKHIEDTG
STPSIGENDLKFTNFNGTENHRENTSRERLWESSSTSHMKIRSQSFNTDTTSGISS
MSSSPRSTVGDATMTDCGSMSTVSTKTKTSHYLTQSHLSLSKSNVSLVP
PGSSHTLPRAQSLKAPSIATIKSLADCNFSTYSSRDAFGYATLKLQOQRMPSLSH
SEALASPAKDLFTDTITMKANSFESRLTPSRPMKALSVALDKEDLLSPINQNTLOR
SSSVRSWSSATVGGSDDIYGLALPVDINDIFQVKDIPYQTKNIPPHDRGARAFAH
DAGLPGSTGGLVKNFSHLLRQMSLTEMNSIHSASLFESEDTGQHEHTDNCCL
YVCBILOQPSNLSAI CSHSDQFDIPEDWCEQTHNPVLEWPFKSGISGCSCL
VSQEGSASTKSTELLGVKTIPTDTPMCRILLRKEVLRLVINLSNVSSTKCHETGLL
TIKEKYPQTFDDICLYSEVSHLLSHCTFLPCRRFTQELFQDVQVFLQMBEABAVLAT
PPKQPIVDTSAB"

ORIGIN

Query Match 76.0%; Score 4300.4; DB 9; Length 8213;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4334; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
Qy 1326 GGCACAAACACATTTCTTCTCATTTACATAGCCATCATTTACATGCTTGGCAACCCCTAAT 1385
Db 1326 GGCACAAACACATTTCTTCTCATTTACATAGCCATCATTTACATGCTTGGCAACCCCTAAT 60
Qy 1386 GAATATGCTGATCTCTTGGATATCCCAAGGAAAGAGACTCGAGCCAGTGCAGCCTT 1445
Db 61 GAATATGCTGATCTCTTGGATATCCCAAGGAAAGAGACTCGAGCCAGTGCAGCCTT 120
Qy 1446 GAACTGTTTAAACCGCTTCCATGAAATGAAGAAACGAGGACCTTAAGCCTTTATAGCTTCA 1505

QY 3666 GAAGATCGTAGCCAAAGTTTCAATACAGACACATACAAACAAGTGGCATAAGTTCAATGAG 3725
DB 2341 GAAGATCGTAGCCAAAGTTTCAATACAGACACATACAAACAAGTGGCATAAGTTCAATGAG 2400
QY 3726 CTCAAGTCCTTTCACGAGACAGTAGTAGTGTAGATGCTACAACTATGAGACACAGACTGTGG 3785
DB 2401 CTCAAGTCCTTTCACGAGACAGTAGTAGTGTAGATGCTACAACTATGAGACACAGACTGTGG 2460
QY 3786 AAGCATGAGTACTGTGGTAAGTACTAAAACTATTAAAGCAAGCCACATATTGAGCGCCACA 3845
DB 2461 AAGCATGAGTACTGTGGTAAGTACTAAAACTATTAAAGCAAGCCACATATTGAGCGCCACA 2520
QY 3846 GTCTAACCATCTGCTCTCTCCAAATCAAAATTCGGTGTCCCTGTGCTCCAGGTTCTTC 3905
DB 2521 GTCTAACCATCTGCTCTCTCCAAATCAAAATTCGGTGTCCCTGTGCTCCAGGTTCTTC 2580
QY 3906 TCATACGCTTCTTAGAGACACAGTCCCTTAAAGACACCCCTCTATTGCTACAAATTAAGA 3965
DB 2581 TCATACGCTTCTTAGAGACACAGTCCCTTAAAGACACCCCTCTATTGCTACAAATTAAGA 2640
QY 3966 TCTAGCAGATTGTAACTTTAGTTACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACT 4025
DB 2641 TCTAGCAGATTGTAACTTTAGTTACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACT 2700
QY 4026 GAAAGACTACAGCAACAAAGATGCATCCATCTCTTATCTCAGTGAAGCTTTGGCAATC 4085
DB 2701 GAAAGACTACAGCAACAAAGATGCATCCATCTCTTATCTCACTCTGAAGCTTTGGCAATC 2760
QY 4086 TCCAGCAAAAGATGTCTATTACTGATPACCATACCAATGAAGCCCAACAGTTTGGATC 4145
DB 2761 TCCAGCAAAAGATGTCTATTACTGATPACCATACCAATGAAGCCCAACAGTTTGGATC 2820
QY 4146 CAGATTAAACCAAGCAGTTTCATGAAGCCTTAAGTTATGTCATCATTTAGATAAAGA 4205
DB 2821 CAGATTAAACCAAGCAGTTTCATGAAGCCTTAAGTTATGTCATCATTTAGATAAAGA 2880
QY 4206 TTTATTGAGTCTTAAATCAAAATACCTGCAACGATCTTCCTCAGTGGGTCATGGT 4265
DB 2881 TTTATTGAGTCTTAAATCAAAATACCTGCAACGATCTTCCTCAGTGGGTCATGGT 2940
QY 4266 GTCCAGTGCCACATATGGGGTTCAGATGATTAATTTGGTCTGCTCTCCCGTGGATAT 4325
DB 2941 GTCCAGTGCCACATATGGGGTTCAGATGATTAATTTGGTCTGCTCTCCCGTGGATAT 3000
QY 4326 AAATGATATATTCAGAGTAAAGGATATTCCTTATTTTTCAGACAAAACATACCAACACA 4385
DB 3001 AAATGATATATTCAGAGTAAAGGATATTCCTTATTTTTCAGACAAAACATACCAACACA 3060
QY 4386 TGATGATCGAGGTGCAAGAGCATTTGCCCATGATGCAAGGAGTCTTCCATCTGGAATGG 4445
DB 3061 TGATGATCGAGGTGCAAGAGCATTTGCCCATGATGCAAGGAGTCTTCCATCTGGAATGG 3120
QY 4446 AGGTCTTGTAATAAAATCTTTTCACTTGTCTACGACAGCATGATGCTTACGGNAATAT 4505
DB 3121 AGGTCTTGTAATAAAATCTTTTCACTTGTCTACGACAGCATGATGCTTACGGNAATAT 3180
QY 4506 GAATTCATCATTCAGATGCTCTCTGTTTATTTAGAAAGTACAGAAGACACTGAGCTACA 4565
DB 3181 GAATTCATCATTCAGATGCTCTCTGTTTATTTAGAAAGTACAGAAGACACTGAGCTACA 3240
QY 4566 GGAAACATACAGATGAATGCTCTTTATTTGCTGCTGTTGAAATCTCGGTTTCCAGCC 4625
DB 3241 GGAAACATACAGATGAATGCTCTTTATTTGCTGCTGTTGAAATCTCGGTTTCCAGCC 3300
QY 4626 CAGCAACCACTGAGTGCATATAGTCAATTCAGACTTTCAGATATTCATATTTCTGA 4685
DB 3301 CAGCAACCACTGAGTGCATATAGTCAATTCAGACTTTCAGATATTCATATTTCTGA 3360
QY 4686 TTGTGTGAGCAGACTATCCATAATCTTTTAGAAGTGTTCCTCTAAGTTTTCGGGGAT 4745
DB 3361 TTGTGTGAGCAGACTATCCATAATCTTTTAGAAGTGTTCCTCTAAGTTTTCGGGGAT 3420
QY 4746 TTCTGGATGCGATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCAACCAAAAGCACAGA 4805

DB 3421 TTCTGGATGCGATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCAACCAAAAGCACAGA 3480
QY 4806 ATTCTTACTAGTGTCTTAAACAATTTCCAGATGATACACCAATGTGCGGTATCTCTTCG 4865
DB 3481 ATTCTTACTAGTGTCTTAAACAATTTCCAGATGATACACCAATGTGCGGTATCTCTTCG 3540
QY 4866 CAAAGAAAGTTCTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAAACTAAATGCTATGA 4925
DB 3541 CAAAGAAAGTTCTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAAACTAAATGCTATGA 3600
QY 4926 GACTGGGCTTTTAAACAATTAAGGAGAGATCTCTCAAAATTTGATGACATATGCTCTTTA 4985
DB 3601 GACTGGGCTTTTAAACAATTAAGGAGAGATCTCTCAAAATTTGATGACATATGCTCTTTA 3660
QY 4986 CTCTGAGGTTTCCCATTTGTGTGCACACTGCACATTCAGACTTCCGCTGCGAGGTTTCAT 5045
DB 3661 CTCTGAGGTTTCCCATTTGTGTGCACACTGCACATTCAGACTTCCGCTGCGAGGTTTCAT 3720
QY 5046 ACAAGAAATTTTCAAGATGTACAGATTTCTCAAAATGCATGAAGAAGCAGAGGCTGTGTT 5105
DB 3721 ACAAGAAATTTTCAAGATGTACAGATTTCTCAAAATGCATGAAGAAGCAGAGGCTGTGTT 3780
QY 5106 GGCAACACCAACCAAGCAACTATAGTTGATACATCTGCTGAATCTCTGACCTCATATTTA 5165
DB 3781 GGCAACACCAACCAAGCAACTATAGTTGATACATCTGCTGAATCTCTGACCTCATATTTA 3840
QY 5166 TGATGGATATAGATACATACATATATATTTTGTGATTTTCTTAAAGCCTCAGA 5225
DB 3841 TGATGGATATAGATACATACATATATATTTTGTGATTTTCTTAAAGCCTCAGA 3900
QY 5226 AAATACGACTGACTAGGACCAAGACAGGAGTATCTTCTGTACACTGTTCCGAGTTAC 5285
DB 3901 AAATACGACTGACTAGGACCAAGACAGGAGTATCTTCTGTACACTGTTCCGAGTTAC 3960
QY 5286 TGGTACATGAACAGTTGGAACTGCTGACTTTTCTTAACCAAAACAACTTCTCTCTCT 5345
DB 3961 TGGTACATGAACAGTTGGAACTGCTGACTTTTCTTAACCAAAACAACTTCTCTCTCT 4020
QY 5346 TGTGTAGCCTTTTGGGGTTCATGATTCATTAACCAAGTTTAAAGAGTTTTCAGTTACCA 5405
DB 4021 TGTGTAGCCTTTTGGGGTTCATGATTCATTAACCAAGTTTTCAGTTACCA 4080
QY 5406 TTGTATGCAAGAGCCAGCACTGAATACCTACATAGGTTTCTTATTTTCTTCAATTTAA 5465
DB 4081 TTGTATGCAAGAGCCAGCACTGAATACCTACATAGGTTTCTTATTTTCTTCAATTTAA 4140
QY 5466 AAGCGTAATGACAGTGGAAACAATAATGGGATATGCAAGACACCTTTCACAAGTTATTC 5525
DB 4141 AAGCGTAATGACAGTGGAAACAATAATGGGATATGCAAGACACCTTTCACAAGTTATTC 4200
QY 5526 TGAATGA-TTTTGGGTAAATATACAGATGCTTGTATGTTAACTTAACCTTGTGGAAGC 5584
DB 4201 TGAATGA-TTTTGGGTAAATATACAGATGCTTGTATGTTAACTTAACCTTGTGGAAGC 4260
QY 5585 AGGAATCAGTGTCTCTAAGGCTGCATCTTATTAACCAATGGGGTGTGCTATAAATGCGC 5644
DB 4261 AGGAATCAGTGTCTCTAAGGCTGCATCTTATTAACCAATGGGGTGTGCTATAAATGCGC 4318
QY 5645 TGGTATTAGAGAGGGAAC 5662
DB 4319 TGGTATTAGAGAGGGAAC 4336

RESULT 6

BD185164 8232 bp DNA linear PAT 17-JUN-2003
LOCUS BD185164
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD185164
VERSION BD185164.1 GI:31877364
KEYWORDS JP 2002345493-A/7.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1	(bases 1 to 8232)							
AUTHORS	Ohara,O., Nagase,T. and Nakajima,D.								
TITLE	Novel genes and proteins encoded by the genes								
JOURNAL	Patent: JP 2002345493-A 7 03-DEC-2002;								
COMMENT	KAZUSA DNA RESEARCH INSTITUTE								
	OS Homo sapiens (human)								
	PN JP 2002345493-A/7								
	PD 03-DEC-2002								
	PF 26-FEB-2002 JP 2002049046								
	PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA								
	PC C12N15/09,C07K14/47,C07K14/54,C12N15/00								
	CC Novel genes and proteins encoded by the genes FH Key								
	Location/Qualifiers								
FT	CDS	(1473).. (3845).							
FEATURES	Location/Qualifiers								
source	1.. 8232								
	/organism="Homo sapiens"								
	/mol_type="genomic DNA"								
ORIGIN	/db_xref="taxon:9606"								
	Query Match 75.4%; Score 4271.4; DB 6; Length 8232;								
	Best Local Similarity 99.5%; Pred. No. 0;								
	Matches 4334; Conservative 0; Mismatches 1; Indels 22; Gaps 4;								
QY	1326	GGCAACACAAATCTCTCATTCACATAGCCATCATTTACATGCTGTGGCAACCTTAAT	1385						
DB	1	GGCAACACAAATCTCTCTCATTCACATAGCCATCATTTACATGCTGTGGCAACCTTAAT	60						
QY	1386	GAATATGGCTGCATCTCTTGTATATCCCAAGGAAAAGAGACTCGAGCCAGTCGAGCCTT	1445						
DB	61	GAATATGGCTGCATCTCTTGTATATCCCAAGGAAAAGAGACTCGAGCCAGTCGAGCCTT	120						
QY	1446	GAATGTTTAAACGGTTCATGAATGAAGAAACGAGGACCTTAAGCCTTATAGTCTTCA	1505						
DB	121	GAATGTTTAAACGGTTCATGAATGAAGAAACGAGGACCTTAAGCCTTATAGTCTTCA	180						
QY	1506	TTTAGACCAATTTATCAGAAAGCAATTGCAACACACACGAAACGGATCAGTATCTCCG	1565						
DB	181	TTTAGACCAATTTATCAGAAAGCAATTGCAACACACGAAACGGATCAGTATCTCCG	240						
QY	1566	AGTTCAGAAAGATATATTTATCTTAAAGATACAGAGAAAGCTCTTTAATTAACCTTAG	1625						
DB	241	AGTTCAGAAAGATATATTTATCTTAAAGATACAGAGAAAGCTCTTTAATTAACCTTAG	300						
QY	1626	AGATAGCAAGTCCTTCAACATAAAGAGAAATCTTGAATGGAATCTTATAGGAC	1685						
DB	301	AGATAGCAAGTCCTTCAACATAAAGAGAAATCTTGAATGGAATCTTATAGGAC	360						
QY	1686	CATTCTTAACTGAGTGGCCAAATCTTAATCTAAGAACTATAAGAGTGAACAGTATACAGGTT	1745						
DB	361	CATTCTTAACTGAGTGGCCAAATCTTAATCTAAGAACTATAAGAGTGAACAGTATACAGGTT	420						
QY	1746	TGTACGAAGACTACTTTTATTTTACAAGCCAGCAGTAAATATATGCCAACCTGGATCT	1805						
DB	421	TGTACGAAGACTACTTTTATTTTACAAGCCAGCAGTAAATATATGCCAACCTGGATCT	480						
QY	1806	GGATTTTGGCAAGGCCAAACAGCTCAAGTGTAGGTTGCCAGTTTACAGAAATTTCTTCT	1865						
DB	481	GGATTTTGGCAAGGCCAAACAGCTCAAGTGTAGGTTGCCAGTTTACAGAAATTTCTTCT	540						
QY	1866	TGAATCTGAAGAGATGGGCAAGGCTACTTTAGAAGATCTAGTAAAGATATTTGTCAGTG	1925						
DB	541	TGAATCTGAAGAGATGGGCAAGGCTACTTTAGAAGATCTAGTAAAGATATTTGTCAGTG	600						
QY	1926	GCTCAATGCTTCAATCTGGAATGAACCCGAAAGAAAGTCTTTCAAAATAATGGTTTATTGAC	1985						
DB	601	GCTCAATGCTTCAATCTGGAATGAACCCGAAAGAAAGTCTTTCAAAATAATGGTTTATTGAC	660						
QY	1986	CACCTTAGTCAACACTACTTTTTTATTTATTTAGGAACACTTTTCTTGGCAACCTTCATGGAGT	2045						

QY 3107 AGTCAACCCAGCTCTAGACATAAATAGTGAAAGTGAATCTGTGCCATCGAGTAGTTCATAT 3166
DB |||||
DB 1801 AGTCAACCCAGCTCTAGACATAAATAGTGAAAGTGAATCTGTGCCATCGAGTAGTTCATAT 1860
QY |||||
DB 3167 TGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCCCTTGATATCAATGAAG 3226
DB |||||
DB 1861 TGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCCCTTGATATCAATGAAG 1920
QY |||||
DB 3227 ATACAGAGCCACACATTTTATGACCGATCTGGACCCATATAAGGATATAAAATTTCAATCCCTT 3286
DB |||||
DB 1921 ATACAGAGCCACACATTTTATGACCGATCTGGACCCATATAAGGATATAAAATTTCAATCCCTT 1980
QY |||||
DB 3287 TCTTTGCTTCTAGTAATACTTGTGAAGAAATCGTATCTTTAAATTCGCTTATCTTGCCTAAACA 3346
DB |||||
DB 1981 TCTTTGCTTCTAGTAATACTTGTGAAGAAATCGTATCTTTAAATTCGCTTATCTTGCCTAAACA 2040
QY |||||
DB 3347 AAAAAACATCTGTAGTACAGTATCCAAAGAGGGAATTTATCATCTGAAAGTAAAGCAA 3406
DB |||||
DB 2041 AAAAAACATCTGTAGTACAGTATCCAAAGAGGGAATTTATCATCTGAAAGTAAAGCAA 2100
QY |||||
DB 3407 GCAACAGCGCAATCAGAACTTACGGAGCCAGTGTGTGATTTTAAATCATAGTATGATTT 3466
DB |||||
DB 2101 GCAACAGCGCAATCAGAACTTACGGAGCCAGTGTGTGATTTTAAATCATAGTATGATTT 2160
QY |||||
DB 3467 TTACACCCATPATCCACTGTACAGAAACATTAACAATTAGAGACTTCAATTTATGGGAATA 3526
DB |||||
DB 2161 TTACACCCATPATCCACTGTACAGAAACATTAACAATTAGAGACTTCAATTTATGGGAATA 2220
QY |||||
DB 3527 AGCATTGGAAGACATGTTAGTACACCAAGCAATGAGAAATGACTTAAATTTACCA 3586
DB |||||
DB 2221 AGCATTGGAAGACATGTTAGTACACCAAGCAATGAGAAATGACTTAAATTTACCA 2280
QY |||||
DB 3587 AGAATTTGTACAGAGAAATCAGAGAAATAACAAGCCAGAGAGGTTAGTAGTAGAA 3646
DB |||||
DB 2281 AGAATTTGTACAGAGAAATCAGAGAAATAACAAGCCAGAGAGGTTAGTAGTAGAA 2340
QY |||||
DB 3647 GTTCAACGAGCTCACATATGAAGATAGCTGACCAAGTTTCAATACAGACACTACAACAA 3706
DB |||||
DB 2341 GTTCAACGAGCTCACATATGAAGATAGCTGACCAAGTTTCAATACAGACACTACAACAA 2400
QY |||||
DB 3707 GTGGCATATGTTCAATGAGCTCAAGTCTTTACAGAGACAGTAGGTGTAGTGTACAA 3766
DB |||||
DB 2401 GTGGCATATGTTCAATGAGCTCAAGTCTTTACAGAGACAGTAGGTGTAGTGTACAA 2460
QY |||||
DB 3767 CTATGACACAGACTGTGGAAGCATGAGTACTCTGTTAGTACTTAAACTTAAAGACAA 3826
DB |||||
DB 2461 CTATGACACAGACTGTGGAAGCATGAGTACTCTGTTAGTACTTAAACTTAAAGACAA 2520
QY |||||
DB 3827 GCCACTATTTGAGCCACAGCTCTAACCATCTGTCTCTCCAAATCAAATTCGGTGTCCC 3886
DB |||||
DB 2521 GCCACTATTTGAGCCACAGCTCTAACCATCTGTCTCTCCAAATCAAATTCGGTGTCCC 2580
QY |||||
DB 3887 TGGTGCTCCAGGTTCTTCTCATACGCTTCTTGAAGAGACAGTCCCTTAAAGCACCTT 3946
DB |||||
DB 2581 TGGTGCTCCAGGTTCTTCTCATACGCTTCTTGAAGAGACAGTCCCTTAAAGCACCTT 2640
QY |||||
DB 3947 CTATTTGCTACAAATTAAGCTTAGCAGATGTTAACTTTAGTTACAGAGTTCTAGAGATG 4006
DB |||||
DB 2641 CTATTTGCTACAAATTAAGCTTAGCAGATGTTAACTTTAGTTACAGAGTTCTAGAGATG 2700
QY |||||
DB 4007 CTTTTCGCTATGCTACACTGAAAGACTACAGCAACAAGAAATGCATCCATCTCTATCTC 4066
DB |||||
DB 2701 CTTTTCGCTATGCTACACTGAAAGACTACAGCAACAAGAAATGCATCCATCTCTATCTC 2760
QY |||||
DB 4067 ACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGA 4126
DB |||||
DB 2761 ACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGA 2820
QY |||||
DB 4127 AGGCCAAAGTTTTGAGTCCAGATTAACCAAGAGCGGTTTCAAGAGCTTTAAGTTATG 4186
DB |||||
DB 2821 AGGCCAAAGTTTTGAGTCCAGATTAACCAAGAGCGGTTTCAAGAGCTTTAAGTTATG 2880

QY 4187 CATCATTTAGATAAAGAGATTTTATTTAGTCTCTATTAATCAAAATACCTGCAACGATCTT 4246
DB |||||
DB 2881 CATCATTTAGATAAAGAGATTTTATTTAGTCTCTATTAATCAAAATACCTGCAACGATCTT 2940
QY |||||
DB 4247 CCTCAGTGGGTCATAGTGTCCAGTGCACATATGGGGGTTCCAGATGATTTACATTTGGTC 4306
DB |||||
DB 2941 CCTCAGTGGGTCATAGTGTCCAGTGCACATATGGGGGTTCCAGATGATTTACATTTGGTC 3000
QY |||||
DB 4307 TTCTCTCTCCCGTGGGATATAAATGATATATTTCCAGGTAAAGGATATTTCCCTATTTTTCAGA 4366
DB |||||
DB 3001 TTCTCTCTCCCGTGGGATATAAATGATATATTTCCAGGTAAAGGATATTTCCCTATTTTTCAGA 3060
QY |||||
DB 4367 CAAAAACATACCAACCATGATGATCGAGGTCAAGAGCAATTTGGCCATGATGACGAG 4426
DB |||||
DB 3061 CAAAAACATACCAACCATGATGATCGAGGTCAAGAGCAATTTGGCCATGATGACGAG 3120
QY |||||
DB 4427 GTCTTCCATCTGAACTGGAGGTCCTTTGTAATAAAATTTCTTTTCACTTGCTACGACAGAGA 4486
DB |||||
DB 3121 GTCTTCCATCTGAACTGGAGGTCCTTTGTAATAAAATTTCTTTTCACTTGCTACGACAGAGA 3180
QY |||||
DB 4487 TGAGTCTTACGGAAATAATGAATTTCAATCCATTCAGATGCTCTCTGTTTTTATAGAAAGTA 4546
DB |||||
DB 3181 TGAGTCTTACGGAAATAATGAATTTCAATCCATTCAGATGCTCTCTGTTTTTATAGAAAGTA 3240
QY |||||
DB 4547 CAGAAGACACTGGACTACAGGAAACATACAGATGATTAACCTGCCCTTTATTTGTGTATTTG 4606
DB |||||
DB 3241 CAGAAGACACTGGACTACAGGAAACATACAGATGATTAACCTGCCCTTTATTTGTGTATTTG 3300
QY |||||
DB 4607 AATTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCATATATGATGATTCATTCAGACTTTC 4666
DB |||||
DB 3301 AATTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCATATATGATGATTCATTCAGACTTTC 3360
QY |||||
DB 4667 AAGATATTTCCATATCTGATTTGGTGAGCAGACTATCCATATCTTTAGAGAGTGGTTC 4726
DB |||||
DB 3361 AAGATATTTCCATATCTGATTTGGTGAGCAGACTATCCATATCTTTAGAGAGTGGTTC 3420
QY |||||
DB 4727 CCTCTAAGTTTTTCGGGGATTTCTGGATGAGTGGGGTGTCTCAAGAGAGGCTCAGCTA 4786
DB |||||
DB 3421 CCTCTAAGTTTTTCGGGGATTTCTGGATGAGTGGGGTGTCTCAAGAGAGGCTCAGCTA 3480
QY |||||
DB 4787 GAGGACCAAAAGCAAGAAATTTGTTACTAGGTGTTTAAACAATTCAGATGATACACCAA 4846
DB |||||
DB 3481 GAGGACCAAAAGCAAGAAATTTGTTACTAGGTGTTTAAACAATTCAGATGATACACCAA 3540
QY |||||
DB 4847 TGTGCGGTATCTCTTCCCAAGAGTTCCTAAGATTTAGTCTAATTTTGGAGTAGTTCAG 4906
DB |||||
DB 3541 TGTGCGGTATCTCTTCCCAAGAGTTCCTAAGATTTAGTCTAATTTTGGAGTAGTTCAG 3600
QY |||||
DB 4907 TTTTCAACTAAATGTCTATGAGACTGGGCTTTTAAACAATTAAGGAGAAAGTATCCTCAAAACAT 4966
DB |||||
DB 3601 TTTTCAACTAAATGTCTATGAGACTGGGCTTTTAAACAATTAAGGAGAAAGTATCCTCAAAACAT 3660
QY |||||
DB 4967 TTGATGACATATGCTTTTACTCTGAGGTTTCCCATTTTGTGTCACACTGCAATTCAGAC 5026
DB |||||
DB 3661 TTGATGACATATGCTTTTACTCTGAGGTTTCCCATTTTGTGTCACACTGCAATTCAGAC 3720
QY |||||
DB 5027 TTTCCGTTGCGAGGTTTCATACAGAAATTTTCAAGATGATGATGATTTTCAAAATGATG 5086
DB |||||
DB 3721 TTTCCGTTGCGAGGTTTCATACAGAAATTTTCAAGATGATGATGATTTTCAAAATGATG 3780
QY |||||
DB 5087 AAGAGCAGAGGCTGTGTTGGCAACACCAAGCAAACTATAGTTGATACATCTGCTG 5146
DB |||||
DB 3781 AAGAGCAGAGGCTGTGTTGGCAACACCAAGCAAACTATAGTTGATACATCTGCTG 3840
QY |||||
DB 5147 AATCTCTGACCTCATATTTTATGATGATATAGATACATCTATATATTTTATTTTGTGG 5206
DB |||||
DB 3841 AATCTCTGACCTCATATTTTATGATGATATAGATACATCTATATATTTTATTTTGTGG 3900
QY |||||
DB 5207 ATTTCTTAAAGCTCAGAAAAATACGATGATGATGATGATGATGATGATGATGATGATG 5266
DB |||||
DB 3901 ATTTCTTAAAGCTCAGAAAAATACGATGATGATGATGATGATGATGATGATGATGATG 3960
QY |||||
DB 5267 TACACTGTTCCGAGTTACTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 5326

```
|||||
3961 TACACTGTTCCGACGTACTGCTACATGAACAGATTGGAACCTGCTGACTTCTTAACCAA 4020
|||||
5327 ACAACTTCCTCTCTCCCTTTGTTGAGCCCTTTGAGGGCTTCATGATTCATTACCACAGTT 5386
|||||
4021 ACAACTTCCTCTCTCCCTTTGTTGAGCCCTTTGAGGGCTTCATGATTCATTACCACAGTT 4080
|||||
5387 TTAAGAGTTTCAGTTTACCACTGTTATGCAAGACCAAGCACTGAATACCCTACATAGGTTTT 5446
|||||
4081 TTAAGAGTTTCAGTTTACCACTGTTATGCAAGACCAAGCACTGAATACCCTACATAGGTTTT 4140
|||||
5447 CTATTTTCTTTCAATTTAAAAGCGTAATGACAGTGGAAACAATAATGGGATATGCAAGAC 5506
|||||
4141 CTATTTTCTTTCAATTTAAAAGCGTAATGACAGTGGAAACAATAATGGGATATGCAAGAC 4200
|||||
5507 ACCCTTCACAGTATTTCTGTAATGA-TTTTAGGTAATAATACAGATGCTTGTGTAT 5565
|||||
4201 ACCCTTCACAGTATTTCTGTAATGATTTTATAGGTAATAATACAGATGCTTGTGTAT 4260
|||||
5566 TAACTAACTTGTGGAAGCAGGAATCAGTGTCTCTAAGGCTGCATCTTATTACCACAATG 5625
|||||
4261 TAACTAACTTGTGGAAGCAGGAATCAGTGTCTCTAAGGCTGCATCTTATTACCACAATG 4320
|||||
5626 GGGTGTGCTATACTGCTGGTATTAGAGGGGAAC 5662
|||||
4321 GGG-TGTGCTATACT-GCTGCTATTAGAGGGGAAC 4355
|||||

RESULT 7
AY497009
LOCUS Mus musculus AVO3 mRNA, complete cds. linear ROD 13-JAN-2004
DEFINITION AY497009
ACCESSION AY497009
VERSION AY497009.1 GI:40737169
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5127)
Loewith, R.J. and Hall, M.N.
Mouse mAVO3
Unpublished
REFERENCE 2 (bases 1 to 5127)
Loewith, R.J. and Hall, M.N.
Direct Submission
TITLE Submitted (08-DEC-2003) Biochemie, University of Basel Biozentrum,
Klingelbergstrasse 50/70, Basel, BS 4056, Switzerland
FEATURES
source
1..5127
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H"
/db_xref="taxon:10090"
/chromosome="15"
/cell_line="C2C12"
1..5127
/note="mAVO3; binds mTOR; similar to yeast AVO3 and
Dictyostelium discoideum planissimo"
/codon_start=1
/product="AVO3"
/protein_id="AAR89074.1"
/db_xref="GI:40737170"
/translation="MAALGRSLKNLVRGNDSEBENPDLITREPSDNLREILQN
VAKUGVSNMRKGLHNNFTKLLCDIGHSBEKJGFNFYEDIIICURLALINEKEVRAA
GLRALYLDQSSILQVLYLKVLYIARCDIQOSNEVERTQALRLVRKMTVNASL
FPSSVANSLIADVNDLQGERDMVQACIAIICELALQNPVVVALRGGLNTILKNVIDC
QLSRINEALIITLHLLNHPKTRQVRADELIERIAPYDFVHRHSPDTAEGOLKED
REARELASKMGIATFRSWAGIINLCKPNSGSIQSLIGVLCIPNMEIRGRLEVIYDI
FRFLPVVDTFIEALLSVDFGRFQDSVRLSDGFVAABAKTILPHRARSFRLMDNYL
ALILSAPFRNGLEGLVEVITNSDDHISVRATILIGELHWNANTILPHSHSHLHCLP
TIMMNASFDIPKPKRLASALNCLNRPHEMKRGPXPYSILHDHIIIOKAIAITHKR
DQYLVRQKDIIFVLKDTTEALLINLSDSQVLQKENLDWNLIGTILKWPVNLRNRYK
```

ORIGIN

Query Match 73.5%; Score 4158.8; DB 10; Length 5127;

Best Local Similarity 88.4%; Pred. No. 0;

Matches 4537; Conservative 0; Mismatches 587; Indels 6; Gaps 2;

```
QY 28 ATGCGCGGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGGCGGAATGAC 87
DB 1 ATGCGCGGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGGCGGAATGAC 60
QY 88 AGCGCGGAGGAGAACGTCCTCGATCTGACCCGAGAACCTTCTGATAAATTAAAGAG 147
DB 61 AGCGCGGAGGAGAACGTCCTCGATCTGACCCGAGAACCTTCTGATAAATTAAAGAG 120
QY 148 ATTCTCMAAATGTGGCAGATTCGAGGAGTATCAATATAGAAAGCTAGGCCATCTG 207
DB 121 ATTTCTCAAAATGTGGCAAAATTCGAAGGAGTATCAAAATATGAGAAAGCTGGGCATCTG 180
QY 208 AATAACTTTTACTAAGCTTCTTTGTGATATGCGCACAGTGAAGAAAAAATCGGGCTTTCAC 267
DB 181 AATACTTCACTAAGCTTCTTTGGATATGCGCATATGCGCATAGTGAAGAAAAAATCGGGCTTTCAT 240
QY 268 TATCAGGATATCATAAATTTGTTGCGGTAGCTTTATTAATGAAGCAAAAGAGTGCGA 327
DB 241 TATGAAGACATCATAAATTTGTTGCGGTAGCTTTATTAATGAAGCAAAAGAGTGAAGA 300
QY 328 GCAGCGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTCCTCCAGAGCTG 387
DB 301 GCAGCGGCTTCGAGCGCTTCGATATCTCATCCAGACTCCAGTATTCCTCCAGAGCTG 360
QY 388 CTAAAAATTGAAAGTGGACTATTTAATAGTAGTGCTGATGATACATACAGAGCAACGAG 447
DB 361 CTAAAACTGAAAGTGGACTATTTAATAGTAGTGCTGATGATGATTCAGCAGAGCAACGAG 420
QY 448 GTAGAGGAGCACAAAGCACTTCGATATAGTTCAGAAAGATGATTAATGTAATGCTTCCCTG 507
DB 421 GTGAGAGGACACAAAGCCCTTCGCTTAGTCAGAAAGATGATTAATGTAATGCTTCCCTG 480
QY 508 TTTCTAGTTCTGTGACCAACTCATTAATTCAGTTGGAATGATGAGCTTCAGAAAGA 567
DB 481 TTTCTAGTTCTGTGCGCAACTCATTAATTCGCTGGAATGATGAGCTTCAGAAAGA 540
QY 568 GACAGATGCTCCGAGCATGATTTGCCATTTCTGTGAATAGCACTTCAGAAATCCAGAG 627
DB 541 GACAGATGCTCCGAGCTTGCATGCCATCATCTGTGAATAGCACTTCAGAAATCCAGAG 600
QY 628 GTGGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAAAATGTGATGATTCGCAATTA 687
DB 601 GTGGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAAAATGTGATGATTCGCAATTA 660
QY 688 AGTCGAATAAATGAGGCCCTTAATTAATCAATTTTGCACCTTCTTAATCATCCAAAGACT 747
DB 661 AGTCGAATAAATGAGGCCCTTCAATTAATCAATTTTGCACCTTCTTAATCATCCAAAGACT 720
```


QY 748 CGG CAGTATGTGCGAGCTGATGTAGAAATTAGAGAGAAATTTTGTAGCACCCTTACTGATTTT 807
DB |||||
QY 721 CGG CAGTATGTGCGAGCTGATGTAGAAATTAGAGAGAAATTTTGTAGCACCCTTACTGATTTT 780
DB |||||
QY 808 CACTACAGACATAGTCCAGATACAGCTGGAAGGACAGCTCAAGAAAGACAGAGAGACGA 867
DB |||||
QY 781 CACTACAGACATAGTCCAGATACAGCTGGAAGGACAGCTCAAGAAAGACAGAGAGACGA 840
DB |||||
QY 868 TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGATCATGGCAGGTTATTATTAAT 927
DB |||||
QY 841 TTTCTAGCCAGTAAATGGGAATCATTCGACATTCGCTCGTGGCAGGTTATTATTAAT 900
DB |||||
QY 928 TTATGTAACCTGGAATTTCTGGGATCCAGTCTCTAATAGGAGTACTTTCGATACCAAAAT 987
DB |||||
QY 901 TTATGTAAGCTGGAATTTCTGGGATACAGTCTCTAATTTGGAGTACTTTCGATACCGAAT 960
DB |||||
QY 988 ATGGAATAGGCGAGCTACTTTGAAGTGTATTTGATATATTTCTCTCTCTCTACCT 1047
DB |||||
QY 961 ATGGAATAGGCGAGGCTCTGCTCGAAGTGTCTCTATGACATCTTTCTCTCTCTCTGCC 1020
DB |||||
QY 1048 GTTCTGATGAGGAGTTTCATAGAAGCACTACTCAGTGTAGATCCAGGAGGTTTCCAAGAC 1107
DB |||||
QY 1021 GTTGTGACGATGAGTTTCATAGAAGCCCTGCTCAGTGTAGATCCAGGAAGTTCCAGGAC 1080
DB |||||
QY 1108 AGTTGGAGGCTTTTCAGATGGCTTTGTGGCAGCTGAGGCAAAAATTAATTTCTCTCATCGT 1167
DB |||||
QY 1081 AGTTGGAGGCTTTCCGATGGCTTTGTGGCAGCTGAGGCAAAAATTAATTTCTCTCTCATCGT 1140
DB |||||
QY 1168 GCCAGATCCAGGCCAGACCTCATGGATTAATTTTGGGACGTGATCTCTCTGCAATTAAT 1227
DB |||||
QY 1141 GCTAGATCCAGGCCAGACCTCATGGACAAATTAATTTGGCTCTGATATCTCTGCAATTAAT 1200
DB |||||
QY 1228 CGTAATGACATTTTGTAGAGGCTTGTAGTGTAGTAAACAACAGTGTATCATATCTCA 1287
DB |||||
QY 1201 CGTAATGACATTTTGTAGAGGCTTGTGTAGTGTAGTAAACAACAGTGTATCATATCTCA 1260
DB |||||
QY 1288 GTTAGAGCTACCATTCCTTTTGTAGGAGCTTTTATCATATGGCAAAACAATTTCTCTCAT 1347
DB |||||
QY 1261 GTGAGGCGGACCATTCCTTTTGTAGGAGCTCTCTGATATGGCAAAACCATCTCTGCCCTCAC 1320
DB |||||
QY 1348 TCATATGACCATCATTTTACATCTGCTTGGCAACCTTAATTAATTTGGTGTGATCTTTGAT 1407
DB |||||
QY 1321 TCGCATAGCCACCATCTACATCTGCTTGGCAACCTCTCATGAACATGGCTGCGCTCTCGAT 1380
DB |||||
QY 1408 ATCCCCAAGGAAGAGACATGCGAGCCAGTGCAGCTTGAACCTTTTAAACGCTTCCAT 1467
DB |||||
QY 1381 ATTCCCAAGGAAGAGAGCTGCGAGCCAGTGCAGCCCTGAACTGTTTGAACCGCTTCCAT 1440
DB |||||
QY 1468 GAAATGAAGAAAACGAGGACCTTAAGCCTTATAGTCTTCAATTTTAGACCAATTAATTCAGAAA 1527
DB |||||
QY 1441 GAGATGAAGAAAGCGAGGACCTTAAGCCTTACAGCCCTTCAATTTAGATCATCATCATTCAGAAA 1500
DB |||||
QY 1528 GCAATTCGAACACACCAAGAAACGGGATCAGATATCTCGAGTTCAGAAAAGATATATTTATC 1587
DB |||||
QY 1501 GCAATCGCAACTCACCAAGCGGATCAGTATCTTCGAGTTCAGAAAAGATATATTTGTT 1560
DB |||||
QY 1588 CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGTCTTCAACAT 1647
DB |||||
QY 1561 CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGTCTTCAACAT 1620
DB |||||
QY 1648 AAAGAGAAATCTTGAATTTGGAATCTTATAGGACCATTTCTTAAGTGGCCAAATGTAT 1707
DB |||||
QY 1621 AAAGAGAAATCTTGAATTTGGAATCTTATAGGACCATTTCTTAAGTGGCCAAATGTAT 1680
DB |||||
QY 1708 AATCTAAGAAACTATAAAGATGAACAGTTACAGAGTTTGTACGAAGTACTTTATTTT 1767
DB |||||
QY 1681 AATCTAAGAAACTATAAAGATGAGCAGTTGACAGGTTTGTGGCAGACTTCTTTACTTT 1740
DB |||||
QY 1768 TACAAGCCAGCAGTAAATTTATATGCCAACTGGATCTGAATTTTGGCAAGGCCAAACAG 1827
DB |||||
QY 1741 TACAAGCCAGCAGCAAACTGTAGCTAGTCTGGATCTGGACTTGGCCAAAGTCCAAAGCAG 1800
DB |||||

QY 1828 CTCA CGTGTAGGTTGCGAGCTTTTACAGAAATTTCTTCTGAAATCTGAAGGATGGGCAA 1887
DB |||||
QY 1801 CTCA CGTGTGTTGTTGCTCAGTTTACAGAAATTTCTGCTCGAGTCTGAAGGATGGGCAA 1860
DB |||||
QY 1888 GGCTACTTAGAAGATCTAGTAAAGGATATTGTTTCAGTGGCTCAATGCTTCAATCTGGAATG 1947
DB |||||
QY 1861 GSATACTTAGAAGATCTCGTGAAGGATATTGTTTCAGTGGCTCAATGCTTCAATCTGGAATG 1920
DB |||||
QY 1948 AAACCCGAAAGAGTCTTTCAAAATTAATGTTTATTTGACCAACCTTACTAGTCAACACTACTTT 2007
DB |||||
QY 1921 AAACCTGAGAGAGCTTTCAGATTAATGGCTTACTGACCACTTACTAGTCAACACTACTTT 1980
DB |||||
QY 2008 TTATTTATTTGGAACACTTTCTTCCACCCTCATGGAGTTTAAATGCTGGAATAATGCACT 2067
DB |||||
QY 1981 TTATTTATCGGAACACTCTCTTGTCTCATCTCATGGAGTCAAAATGCTGGAAGAGTGCAGT 2040
DB |||||
QY 2068 GTATTTTCAGTGTCTCTTAATCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2127
DB |||||
QY 2041 GTCTTTTCAGTGTCTCTTAACTTTGTTCTTGAATAACCAAGATCATCTGATAAAACTC 2100
DB |||||
QY 2128 ACTGTTTCTAGCTTGGACTATAGCAGATGGAATTTGGCTAGAGTCACTCTTCCAAAT 2187
DB |||||
QY 2101 ACTGTTTCCAGCTTAGACTATAGCAGATGGAATTTAGCAGAGTCACTCTTCCAAAT 2160
DB |||||
QY 2188 TTAACGTGAGTACTGATGCTCGAGCTCTATGCAACAAAATTTAAGGGTATTAATG 2247
DB |||||
QY 2161 CTCACGCGCAGCAGATGCTTCAGACTCTATGCAACAAAATTTAAGGGTATTAATG 2220
DB |||||
QY 2248 AGAGCTTAATGTTGAATCTTTAATTTGGGGAATTTGAGTTGTTAGTGACCCAGCTACAT 2307
DB |||||
QY 2221 AGAGCTTAACGTTGAATCTTTCAACAAATTTGGGGAATTTGAGTTACTAGTAACCTAGCTACAT 2280
DB |||||
QY 2308 GATAAAAAACAAACGATTTCTCTGAAGCTTTGATATCTCGATGAAGCATGTGGAAGAC 2367
DB |||||
QY 2281 GATAAAAAACAAACAAATTTCTCTGAAGCTGATATCTCGATGAAGCATGTGGAAGAC 2340
DB |||||
QY 2368 AAGGCCAATCTTCATGCTCTCATTTAGATGAACACCGGTTATCCCACTTGGAGACAAG 2427
DB |||||
QY 2341 AAGGCCAATCTTCATGCTCTTATTCAGATGAACCGGCTTATCCCACTTGGAGACAAG 2400
DB |||||
QY 2428 GGTTCGTTCTCTGCTGAGATTTCTCTCATTTCCAAAGGATTTTCTTACTGTAATGAA 2487
DB |||||
QY 2401 GGCTTGTCTTCTACTGCTGAGATTTCTTCCATTTCCGAAAGGATTTTCTTACTGTAATGAA 2460
DB |||||
QY 2488 AGAGGTTATGTAGCAAAACAATTTGGAAAAGTGCACAGGGAATACAACTCCAAATATGTT 2547
DB |||||
QY 2461 AGGGGTTATGTTCGCAAAACAGTTTGGAAAAGTGCACAGGGAATATTAATTTCAAAGTATGTT 2520
DB |||||
QY 2548 GACTTGAATTTGAGGAACAACCTCAATGAAGCACTTACTACTTACCGGAGCCCTGTTGATGTT 2607
DB |||||
QY 2521 GACTTGAATTTGAGGAACAACCTTATGAAGCCCTCACTACATACAGGAAGCCCTATTGATGTT 2580
DB |||||
QY 2608 GATACTATTTGCTGCGAGTAAACAAAGATTTACAGGCTCTCACTGCTTACCTGCTCTATA 2667
DB |||||
QY 2581 GATACTATTTGCTGCGAGTAAACAAAGATTTACAGGCTCTCACTGCTTACCTGCTCTATA 2640
DB |||||
QY 2668 CACTTTATTTGGAACAACCTAGTACACCAATAAAACAGGCTGCCATTTGTTGGAAGTACAGAT 2727
DB |||||
QY 2641 CACTTTATTTGGAACAACCTAGTACACCAATAAAACAGGCTGCCATTTTACTGGAAGTACAGAT 2700
DB |||||
QY 2728 ATTTATTTACAGAACTCTGCTGATATTTGCTGATACCAAGATTTGGAATTTGGAAGTACAGAT 2787
DB |||||
QY 2701 ATCATCACAGAGCTCTGTCACATGTTCCACCTCCAGACTTGGACAAATGGGAAGACAT 2760
DB |||||
QY 2788 AAAAACTGAAGACATCTCTTTTGGGCTTGGGAAATATCGGCTCATCAATTTGGGCTCTC 2847
DB |||||
QY 2761 AAGAGCTGAAGACCTCTCTTTTGGGCTTGGGAAATATTTGGCTCATCGAATTTGGGCTCT 2820
DB |||||
QY 2848 AATTTCTACAGGAAGAAACGATTTCCAGATATATCTAATAAATTTCCAAACAGTGTGAA 2907
DB |||||
QY 2821 AATTTCTACAGGAAGAAATGATATCCAGATATATCTAATAAATTTGCAAAACAGTGTGAG 2880
DB |||||
QY 2908 GTTCTTTTCCATCAGAGGAGCTGTGATATATGTTGCTTGGGCTCATAGCTTAAACCAACAA 2967
DB |||||

2881 GTCCCTTCCATCCAGGGACCTGTGTATATGTTCTTGGGCTCATAGCTAAACCAAGCAA 2940
2968 GGCTGTGATATTCATAAATGTCAAACTGGGATGCTGTGAGGATAGTCGCAAAATCTGT 3027
2941 GGCTGTGATATTTGAAATGTCACTACCTGGGATTTCTGTAAGGCACATAGTCGCAAGCATCTA 3000
3028 TGGCCAGTGTGTCAGATGATGGAAACAACTCTGTAAATGAACTTTCATCTATCCCAAGC 3087
3001 TGGCCAGTGTGTCAGATGATGGGAGAGCTCTGCAATGAGCTCTCATCTGTGCCAGC 3060
3088 ACTCTAAGTTTGAACCTCGAGTCAACCCAGCTCTAGACATAAATAGTGAAGTGAATCTGTG 3147
3061 ACCCTCAGCTTGAACCTCTGAGTCGACAGCTCTAGACATAAAGTGAAGTGAATCTGTG 3120
3148 CCATCCAGTATGTTTCATATGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3207
3121 CCATCAAGCATGTTTCATGCTGGAGGACGACCGCTTCCGAGCAGCCTCCACCAAGCAGCATTT 3180
3208 TTCCCTTGATATCAATGAAGATACAGAGCCAAACATTTTATGACCGATCTGGACCCATAAAG 3267
3181 TTCCCTTGATATCAATGAAGATGCTGAGCCAGCAATTTTATGACCGACCTGGACCCATAAAG 3240
3268 GATAAATAATCATCTCCCTTCTTCTGCTCTAGTAAATCTGTGAAGAACTGTATCTTAAT 3327
3241 GATAAATAATCATCTCCCTTCTTCTGCTCTAGTAAATCTGTGAAGAACTGTATCTTAAT 3300
3328 TCGCTTACTTTCGCTTAACAAAACATCGTAGTAGCAGTATCCAAAGAGGAGGAAATTA 3387
3301 TCACCTTACTTTCGCTTAAATAAACACCGCAGTAGCAGTATCCAAAGAGGAGGAAACCTG 3360
3388 TCATCTGAAAGTAAAGCAAGCAACAGCGGAATCAGAAACATTTACGAGGCCAGTGTGAT 3447
3361 TCATCTGAAATTAAGACAGCAACCGCGAATCAGAAACATTTACGAGGCCAGTGTGAT 3420
3448 TTTAATCATAGTATGATTTTACCCATATCCATATCACTGTACAGAAACATTAACAATTAGAG 3507
3421 TTGAATCAGAGGAGACTTTCACA--AGCTCTCAGCACAGAAATCATTAAGCTAGAG 3477
3508 ACTTCAATTTATCGGGAATAAGCACATGGAAGCACATGGTAGTACACCAAGCATTCGAGNA 3567
3478 CTTTCTCTCGTTGGGAATAAGCACCTTGAAGACGCGGCGAGCAGCTCCAAAGTATGGAGAG 3537
3568 AATGACTTAAATTCACCAAGAAATTTGGTACAGAGAAATCAGAGAAATAACAGCCGA 3627
3538 AATGACTTCAAAATTCACCAAGAAATTTGGTACCGAGACTCAGAGAAATAACAGCCGA 3597
3628 GAGAGTTAGTATGATAAAGTTCAAGAGCTCACATATGAAGATACGTAGCCAAAGTTTC 3687
3598 GAAAGGTTGGTTGTGGAGGCTCAGCAAGCTCCACATCAAGATCCGACGCGAGAGCTTT 3657
3688 AATACAGACACTACAAAGTGGCATAGTTCAATGAGCTCAAGTCTTCAAGGAGACA 3747
3658 AACAGGACACTACAGAGTGGCATAGTTCCATGAGCTCGAGCCCTTCCGAGAGACA 3717
3748 GTAGGTGTAGTGTACAACTATGGACACAGACTGTGGAAGCATAGTACTGTGTGAAGT 3807
3718 GTCCGTGTGACCCACAGCTATGACACGAGCTGTGGAAGTCTGAGTACCGTGTGAGT 3777
3808 ACTAATACTATTAAAGACCACTATTGAGCGCACAGCTCTAACCATCTGTCTCTCTCC 3867
3778 ACTAATACTATTAAAGACCACTATTGACACCACTATTGACACCACTCTGTCTCTCTCC 3837
3868 AATCAAAATTCGGTGTCCCTGTGTGCTCCAGGTCTTCTCATACGCTTCTTAGAGAGACA 3927
3838 AATCGAAGTCAAGTGTCCCTGTGTGCAACAGGTTCTTCTCACACCTTCTTAGAGAGCC 3897
3928 CAGTCCCTTAAAGCACCTCTATTGCTCAATTAAGGCTTAGCAGATTTGAATTTAGT 3987
3898 CAGTCCCTTAAAGCGCCCTCCATTGCAACAAATAAAGTCTAGCAGATTGCAACTTCAGC 3957
3988 TACACAGTTCTTAGAGATGCTTTGGCTATGCTACACTGMAAAGACTACAGCAACAAAGA 4047

3958 TACACAAGCTCTAGAGATGCGCTTTGGCTATGCTCATTTGAAAAGGTTACAGCAGCAAGA 4017
4048 ATGCAATCCATCTTATATCTCATCTGAAGCTTTGGCAATCTCCAGCAAAAAGATGTCTATTT 4107
4018 ATGCAACCGCTCTTGTCTCACTCGAAGCTTTGGCGTCCGCGCAAAAAGATGTCTTTTT 4077
4108 ACTGATACCATCACCATGAAGSCCAAGCTTTGAGTCCAGATTAACCAACCAAGCAGGTTTC 4167
4078 ACCGACACATCACCATGAAGSCCAAGCTTTGAAATCCAGACTGACGCAAGCAGGTTT 4137
4168 ATGAAAGCTTTAAGTTATGCTATCATTTAGATAAAGAAAGATTTTATGAGTCTTATTAATCAA 4227
4138 ATGAAAGCTTTAAGTTATGCTATCATTTAGATAAAGAAAGATTTTACTAAGTCTTATTAATCAC 4197
4228 AATACCCCTGCAACGATCTTCTCAGTGGCTCCATGTTGTCAGTGCACATATGGGGGT 4287
4198 AATACCCCTCCAGCGCTCTCTCAGTTAGGTCTATGTTGTCAGTGCACATATGGAGT 4257
4288 TCAGATGATTAATGTTGTTGCTCTCCCGTGGATATAAATGATATATTCAGGTAAAG 4347
4258 TCCGATGACTACATTTGGACTTGTCTCTCAGTAGACATCAATGACATCTTCCAGATAAAG 4317
4348 GATATTCCTATTTTCAGACAAAACATACACACATGATGATCGAGGTGCAAGACA 4407
4318 GATGTTCTTATTTTCAATCAAGCAGCTGCTCTCCACAGATGACCGAGGTGCAAGAATG 4377
4408 TTTGCCCATGATGACGAGGTTCTTCCATCTGGAATGAGGTCTGTGTAATAAATTTCTTTT 4467
4378 TTTTCCCATGATGAGGAGGTTCTTTCATCTGGAGCGGTGCTCTGTAAGAATCTTTT 4437
4468 CACTTGCTTACGACAGCAGATGAGTCTTACCGAAATTAATGAATTCATTCATTCAGATGCC 4527
4438 CACTTGCTCGGCGCAGCAGATGAGTCTCAGAAATTAATGAATCTGTCCATTCGAGCGCC 4497
4528 TCTCTGTTTTTGAAGATPACAGAGACACTGGAATACAGGAAACATACAGATGATAACTGC 4587
4498 TCCCTGTTCTTTAGAGACACAGAAAGACACTGGGCTGCGAGGAGCACACGATGACAAATTCG 4557
4588 CTTTATTTGTTGCTGTATTTGAAATTTCTGGGTTTCCAGCCCAACCACTGAGTGCAATA 4647
4558 CTTTACTGCGTCTGATTTGAATCTCTGGGCTTCCAGCCCACTGAGTAAAGTTCCATC 4617
4648 TGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATTTGTTGAGCAGACTATCCAT 4707
4618 TGCAGTCACTCCGATCTTCAAGATATTCGCTACTCCGATTTGTTGCGAGCAGACGATCCAC 4677
4708 AATCCTTTGAAGTGGTTTCCCTCTAAGTTTTCGGGGAATTTCTGAGTGCAGTGTGGGGTG 4767
4678 AACCCCTTTGAAGTGGTTTCCCTCTAAGTTTCTCGGGGATTTCTGGCTGCGAGGATGGCGCA 4737
4768 TCTC--AAGAAAGCTCAGCTAGCAGCACCAAAAGCACAGAAATTTGTTACTAGGTGTTAAA 4824
4738 TCTCAGGAAGAGGCTCAGCAGCAGCACCAAAAGCACAGAAATTTGTTACTGGGTGTTAAA 4797
4825 ACAATTCAGATGATACACCAATGTGCGGTATCTCTCTCCAAAGAAAGTTCTAAGATTA 4884
4798 ACCATTCAGATGACACACCCATGTGCGGATCTCTGCTTCTGTAAGAAAGTCTTAAGATTA 4857
4885 GTCAATTAATTTGAGTAGTTTCAAGTTTCAACTAAATGTGATGAGACTGGGCTTTTAAACAAAT 4944
4858 GTTGTCAATTTGAGTAGTTTCAAGTTTCCCAAGTGTGTCAGAGACCGGGCTTCTGACCAAT 4917
4945 AAGGAGAGATCTCTCAAAATTTGATGACATATGCTCTTACTCTGAGGTTTCCCATTTG 5004
4918 AAGAGAGATATCTCTCAAAATTTGATGACATATGCTCTTATTTCTGAGTTTCTCATCTG 4977
5005 CTGTCACTGCAATTCAGACTTCCGTTGCGGAGTTTCATACAGAAATTTATTTCAAGAT 5064
4978 TTGTCACTGCAATTTAGACTTTCAGTGTGCGGGTTTCATACAGAGATTTATTTCCAGAT 5037
5065 GTACAGTTTCTACAAATGCAATGAAGAGCAGAGGCTGTGTGGCAACACCCCAAGCAA 5124
5038 GTACAGTTTCTACAAATGCAATGAAGAGCAGAGGCTGTGTGGCAATTTCCCAATTCAG 5097

[illegible]

3006 GAGGCATAGTCGCAAAACATCTGTGGCCAGTGGTTCCAGATGATGTGGAACAACTCTGTAA 3065
Db |
1681 GAGGCATAGTCGCAAAACATCTGTGGCCAGTGGTTCCAGATGATGTGGAACAACTCTGTAA 1740
QY |
3066 TGAACCTTTCACTATCCCAAGCACTCTAAGTTTGAACCTCGAGTCAACAGACTCTAGACA 3125
Db |
1741 TGAACCTTTCACTATCCCAAGCACTCTAAGTTTGAACCTCGAGTCAACAGACTCTAGACA 1800
QY |
3126 TAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTGGAGGATGACCGGTTTGG 3185
Db |
1801 TAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATATTGGAGGATGACCGGTTTGG 1860
QY |
3186 CAGCAGCTCTACTAGTACATTTTCTTGATATCAATGAAGATACAGAGCCAACTTTTA 3245
Db |
1861 CAGCAGCTCTACTAGTACATTTTCTTGATATCAATGAAGATACAGAGCCAACTTTTA 1920
QY |
3246 TGACCGATCTGGACCCATAAAGGATAAAAATTCATTTCCCTTTCTTTGCTTCTAGTAAACT 3305
Db |
1921 TGACCGATCTGGACCCATAAAGGATAAAAATTCATTTCCCTTTCTTTGCTTCTAGTAAACT 1980
QY |
3306 TGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAACTTAACTTAACTTAACTTAACT 3365
Db |
1981 TGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAACTTAACTTAACTTAACTTAACT 2040
QY |
3366 TGATCCAAAGAGGAGGAAATTTATCATCTGAAAGTAAAGCAAGCAAGGCGCAATCAGAAC 3425
Db |
2041 TGATCCAAAGAGGAGGAAATTTATCATCTGAAAGTAAAGCAAGCAAGGCGCAATCAGAAC 2100
QY |
3426 ACTTAGCGAGCCAGTGTGATTTTAAATCATAGTATGATTTTAAACCCATATPCCACTGT 3485
Db |
2101 ACTTAGCGAGCCAGTGTGATTTTAAATCATAGTATGATTTTAAACCCATATPCCACTGT 2160
QY |
3486 ACAGAAAAATTTACAAATAGAGACTTCAATTTATGGGGAATAGCAATTTGAGACACTGG 3545
Db |
2161 ACAGAAAAATTTACAAATAGAGACTTCAATTTATGGGGAATAGCAATTTGAGACACTGG 2220
QY |
3546 TAGTACACAAAGCATTTGGAGAAATGACATTTAAATTTCAACAAATTTTGGTACAGAGAA 3605
Db |
2221 TAGTACACAAAGCATTTGGAGAAATGACATTTAAATTTCAACAAATTTTGGTACAGAGAA 2280
QY |
3606 TCACAGAGAAATPACAAGCCGAGAGAGGTTAGTGTAGTAAAGTTCAACAGACTCACATAT 3665
Db |
2281 TCACAGAGAAATPACAAGCCGAGAGAGGTTAGTGTAGTAAAGTTCAACAGACTCACATAT 2340
QY |
3666 GAAGATACGTAGCCAAAGTTTCAATACAGCACTACAACAGTGGCATTAAGTTCAATGAG 3725
Db |
2341 GAAGATACGTAGCCAAAGTTTCAATACAGCACTACAACAGTGGCATTAAGTTTCAATGAG 2400
QY |
3726 CTCAGTCTCTTCACAGAGACAGTAGGTGTAGATGCTACAACCTATGGACACAGACTGTGG 3785
Db |
2401 CTCAGTCTCTTCACAGAGACAGTAGGTGTAGATGCTACAACCTATGGACACAGACTGTGG 2460
QY |
3786 AAGCATGAGTACTGTGGTAAGTACTAAAACTATTAAGACAAGCCACTATTTGACGCCACA 3845
Db |
2461 AAGCATGAGTACTGTGGTAAGTACTAAAACTATTAAGACAAGCCACTATTTGACGCCACA 2520
QY |
3846 GTCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTCCCTCCAGGTTCTTC 3905
Db |
2521 GTCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTCCCTCCAGGTTCTTC 2580
QY |
3906 TCATACGCTTCTTAGAAGACAGCAGTCCCTTAAAGCACCCCTCTATTGTCTACAAATTAAGA 3965
Db |
2581 TCATACGCTTCTTAGAAGACAGCAGTCCCTTAAAGCACCCCTCTATTGTCTACAAATTAAGA 2640
QY |
3966 TCTAGCAGATTTGAACTTTAGTTTACAAGTCTCTAGAGATGCTTTTGGCTATGCTACACT 4025
Db |
2641 TCTAGCAGATTTGAACTTTAGTTTACAAGTCTCTAGAGATGCTTTTGGCTATGCTACACT 2700
QY |
4026 GAAAGACTACAGCAACAAGATGATCCATCCCTTATCTCACTCTGAGCTTTGGCATC 4085
Db |
2701 GAAAGACTACAGCAACAAGATGATCCATCCCTTATCTCACTCTGAGCTTTGGCATC 2760
QY |
4086 TCCAGCAAAAGATGTGCTATTTTACTGATACCATCACCATGAAGGCCAACACAGTTTGTGAGTC 4145

2761 TCCAGCAAAAGATGTGCTATTTTACTGATACCATCACCATGAAGGCCAACAGTTTTCAGTC 2820
QY |
4146 CAGATTAAACACCAAGCAGGTTTCATGAAGCCCTTAAGTTATGATCATTAATAGATAAAGA 4205
Db |
2821 CAGATTAAACCAAGCAGGTTTCATGAAGCCCTTAAGTTATGATCATTAATAGATAAAGA 2880
QY |
4206 TTTATTGAGTCTTATTAAATCAAAATACCTCGCAACGATCTTCTCAGTGGGTCCTATGT 4265
Db |
2881 TTTATTGAGTCTTATTAAATCAAAATACCTCGCAACGATCTTCTCAGTGGGTCCTATGT 2940
QY |
4266 GTCCAGTGCCACATATGGGGGTTTCAGATGATTACATTTGCTTCTCTCCCGTGGATAT 4325
Db |
2941 GTCCAGTGCCACATATGGGGGTTTCAGATGATTACATTTGCTTCTCTCCCGTGGATAT 3000
QY |
4326 AAATGATATATTCCAGGTHAAGGATATTTCCCTATTATTTTCAGACAAAAAATACACCA 4385
Db |
3001 AAATGATATATTCCAGGTHAAGGATATTTCCCTATTATTTTCAGACAAAAAATACACCA 3060
QY |
4386 TGATGATCGAGTGCAGAGACATTTGCCCATGATGCAGGAGGTCTTCCATCTGGAACCTGG 4445
Db |
3061 TGATGATCGAGTGCAGAGACATTTGCCCATGATGCAGGAGGTCTTCCATCTGGAACCTGG 3120
QY |
4446 AGGTCTTGTAAAAAATTTCTTTTCACTGTCTACGACGACGATCAGTCTTTACCGAAATA 4505
Db |
3121 AGGTCTTGTAAAAAATTTCTTTTCACTGTCTACGACGACGATCAGTCTTTACCGAAATA 3180
QY |
4506 GAATTCATATCCATTCAGATGCCTCTCTGTTTATAGAAATGACAGACACATGGAATAC 4565
Db |
3181 GAATTCATATCCATTCAGATGCCTCTCTGTTTATAGAAATGACAGACACATGGAATAC 3240
QY |
4566 GGAACATACAGATGATACTGCTTTATTTGCTGTATTTGAAATTTCTGGGTTTCCAGCC 4625
Db |
3241 GGAACATACAGATGATACTGCTTTATTTGCTGTATTTGAAATTTCTGGGTTTCCAGCC 3300
QY |
4626 CAGCAACCAACTGAGTGCAATATGATGTCATTCAGACTTTTCAAGATATTTCCATATTCGA 4685
Db |
3301 CAGCAACCAACTGAGTGCAATATGATGTCATTCAGACTTTTCAAGATATTTCCATATTCGA 3360
QY |
4686 TTGGTGTGAGACAGACTATCCATAATCTTTAGAAAGTGGTTCCTCTAAGTTTTCGGGAT 4745
Db |
3361 TTGGTGTGAGACAGACTATCCATAATCTTTAGAAAGTGGTTCCTCTAAGTTTTCGGGAT 3420
QY |
4746 TTCTGATGTCAGTGTGGGTCCTCAAGAGGCTCAGCTAGCAGCACCAAAAGACACAGA 4805
Db |
3421 TTCTGATGTCAGTGTGGGTCCTCAAGAGGCTCAGCTAGCAGCACCAAAAGACACAGA 3480
QY |
4806 ATTGTTACTAGTGTGTTAAAAACAATTCAGATGATACACCAATGTGCGGTATATCTCTCG 4865
Db |
3481 ATTGTTACTAGTGTGTTAAAAACAATTCAGATGATACACCAATGTGCGGTATATCTCTCG 3540
QY |
4866 CAAAGAGTTCTAAGATTAGTCAATTAATTTGAGTGTTCAGTTTCACTAAATGTCATGA 4925
Db |
3541 CAAAGAGTTCTAAGATTAGTCAATTAATTTGAGTGTTCAGTTTCACTAAATGTCATGA 3600
QY |
4926 GACTGGCTTTTAAACAAATTAAGGAGATATCCTCAACATTTGATGACATATGCCTTTA 4985
Db |
3601 GACTGGCTTTTAAACAAATTAAGGAGATATCCTCAACATTTGATGACATATGCCTTTA 3660
QY |
4986 CTCTGAGGTTTCCCATTTGCTGTCACTGTCACTGTCACTGTCCGAGGTTTCAT 5045
Db |
3661 CTCTGAGGTTTCCCATTTGCTGTCACTGTCACTGTCCGAGGTTTCAT 3720
QY |
5046 ACAAGAAATTTTCAAGATGATCAGTTTCTCAAAATGATGAAGACAGAGCTGTGTT 5105
Db |
3721 ACAAGAAATTTTCAAGATGATCAGTTTCTCAAAATGATGAAGACAGAGCTGTGTT 3780
QY |
5106 GCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTGAATCCTGA 5154
Db |
3781 GCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTGAATCCTGA 3829

BC051729
LOCUS Homo sapiens hypothetical protein MGC39830, mRNA PRI 27-JAN-2004
DEFINITION IMAGE:5787163), partial cds.
ACCESSION BC051729
VERSION BC051729.1 GI:30704351
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5440)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,
Diatchenko, L., Maruina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., Skalska, U., Smailus, D.E.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 5440)
Straussberg, R.
Direct Submission
Submitted (01-MAY-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 88 Row: p Column: 2.

FEATURES

source
1..5440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IWAGR:5787163"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH_MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1..5440
/gene="MGC39830"
/note="synonym: KIAA1999"
/db_xref="LocusID:253260"
<1..3156
/gene="MGC39830"

gene

CDS

/codon_start=1
/product="MGC39830 protein"
/protein_id="AAHS1729.1"
/db_xref="GI:30704352"
/db_xref="LocusID:253260"
/translation="HYFLFIGTISCHPHGVMLEKCVFQCLLNLCSLNQNDHLLKLT
VSSLDYSDGLARVILSLTAATADACRLATYKHLRLVLRANVEFFNNWGIELVTLQ
HDKNTTSSEALDLEACEDKANLHAIQMKPALSHLGKGLLLLRFLSLPKGFSY
LNERGYAKOLEKWHREYKFDVQIIEQLEALTYKRPVGVGNYVRRSNORLGRPH
VYLIHLYGQVHHKTHCHLEVDNIITELCRNVTPDLKWEIEKKLKASLWALGNI
GSSNWGLNLQBEENVIPDLKLAQCEVLSIRGTCVYVLGLIAKTKGCDILKCHND
AVHRSKHLPVVPDVEQCNELSLIPSTLSNSESSTSSRHNSSESSEVSSMFLLE
DRFGSSSTFFLDINEDTEPTFYDRSGPIKDNKSPFFFPASSKLKRNRIILSLPK
KHIEDTSPSGKLSSESKTNRIRITLTSPSVDPNHSDDPTPISTVQTKLQLETSFMG
NKHIEDTSPSGKLSSESKTNRIRITLTSPSVDPNHSDDPTPISTVQTKLQLETSFMG
TTSVGLSSMSSSPSRVGVDAVTMDTCGSMSTVVSTKTKTSHVLTQSNHLSLSK
SNSVLPVPOSSHLPRAQSLKAPSIATIKSLADCNFSTSSRDARFATLKKLQQQ
RMPUSLHSEALAPAKVLTFTITMKNSESRUTPSRPMKALSAYSLDKDLSP
INQNTLQSSSVSMSSATYGGSDYIGLALPVDINDIFQVKDIPYFQTKNIIPHD
RGARAFADAGGLPSGTGLVKNFSHLLRQOMSLTEIMNSIHSDASIFLSESTDTGLQ
BHTDDNCLYCVCIILGFPQSNOLSAICSHSDFDIPIYSDWCEQIHNPLEVVPKFS
GISCSGVSOEGSASTKSTELLGVKTIPTDTPMCRILLRKEVLRLVNLSSVST
KCHETGLLTIKKYPQTFDDICLYSEVSHLLSHCTFLPCRRFIQELFQDVQFLQME
EAEAVLATPPRPPIVDTSAES"

ORIGIN

Query Match	Score	3622,6;	DB 9;	Length	5440;
Best Local Similarity	99.8%;	Pred. No. 0;			
Matches	3659;	Conservative	0;	Mismatches	4;
				Indels	3;
				Gaps	3;
Qy 1999	CACTACTTTTATTTATTTGGAACACATTTCTTTCGCCACCTCATGGAGTAAATGCTCGAA	2058			
Db 1	CACTACTTTTATTTATTTGGAACACATTTCTTTCGCCACCTCATGGAGTAAATGCTCGAA	60			
Qy 2059	AAATGCGAGTGATTTGAGTGCTCTCTTAATCTTTTGTCTTGAATAACCAAGATCATCTG	2118			
Db 61	AAATGCGAGTGATTTGAGTGCTCTCTTAATCTTTTGTCTTGAATAACCAAGATCATCTG	120			
Qy 2119	CTAAACTTCTACTGTTTCTAGCTTGGCATATAGCAGAGATGGATTGGCTAGAGTCACTCTT	2178			
Db 121	CTAAACTTCTACTGTTTCTAGCTTGGCATATAGCAGAGATGGATTGGCTAGAGTCACTCTT	180			
Qy 2179	TCCAAAATTTTAACTGCAGCTACTGATGCTCTGAGACTCTATGCAACAAAACATTTAAGG	2238			
Db 181	TCCAAAATTTTAACTGCAGCTACTGATGCTCTGAGACTCTATGCAACAAAACATTTAAGG	240			
Qy 2239	GTATTTATGAGAGCTAATTTGTAATCTTTTAATAATTTGGGGAATTTAGTTTGTAGTACC	2298			
Db 241	GTATTTATGAGAGCTAATTTGTAATCTTTTAATAATTTGGGGAATTTAGTTTGTAGTACC	300			
Qy 2299	CAGCTACATGATAAAACAAACGATTTCTCTGAGCTCTTGTATATCTCGATGAAGCA	2358			
Db 301	CAGCTACATGATAAAACAAACGATTTCTCTGAGCTCTTGTATATCTCGATGAAGCA	360			
Qy 2359	TGTGAAGACAAAGGCCAATCTTCAATGCTCTCATTCAGATGAACACGAGCTTATCCACCTT	2418			
Db 361	TGTGAAGACAAAGGCCAATCTTCAATGCTCTCATTCAGATGAACACGAGCTTATCCACCTT	420			
Qy 2419	GGAGACAAAGGTTTGTCTTCTGCTGAGATTTCTCTCCATTCCAAAAGGATTTTCTAT	2478			
Db 421	GGAGACAAAGGTTTGTCTTCTGCTGAGATTTCTCTCCATTCCAAAAGGATTTTCTAT	480			
Qy 2479	CTGAATGAAGAGGTTTATGTAGCAAAACAAATTTGGAAGTGGGACAGGGGAATCAACTCC	2538			
Db 481	CTGAATGAAGAGGTTTATGTAGCAAAACAAATTTGGAAGTGGGACAGGGGAATCAACTCC	540			
Qy 2539	AAATATGTTGACTTGTATGAGGAAACAACTCAATGAAGACCTTACTTACTTACCCGAGCCT	2598			
Db 541	AAATATGTTGACTTGTATGAGGAAACAACTCAATGAAGACCTTACTTACTTACCCGAGCCT	600			
Qy 2599	GTTGATGGTGATAAATATGTTCTCGAGAGTAAACAAAGATTACAGCGTCTCTCAGCTAC	2658			
Db 601	GTTGATGGTGATAAATATGTTCTCGAGAGTAAACAAAGATTACAGCGTCTCTCAGCTAC	660			

QY	2659	CTGCTATACACCTTTATGGACAACCTAGTACACCATAAAAACAGGCTGCCATTTGTTGGAA	2718
DB	661	CTGCTATACACCTTTATGGACAACCTAGTACACCATAAAAACAGGCTGCCATTTGTTGGAA	720
QY	2719	GTACAGAAATATATTACAGAACTCTGTCGTAAATGTTTCGTACACAGAAATTTGGATTAAGTGG	2778
DB	721	GTACAGAAATATATTATACAGAACTCTGTCGTAAATGTTTCGTACACAGAAATTTGGATTAAGTGG	780
QY	2779	GAAGAAATTAAGAACTGAAGCAATCTTTGGGCCCTTGGGAATATCGGCTCATCAAAAT	2838
DB	781	GAAGAAATTAAGAACTGAAGCAATCTTTGGGCCCTTGGGAATATCGGCTCATCAAAAT	840
QY	2839	TGGGGTCTCAATTTTGCTACAGGAAGAAAACGTGATTCAGATATACATAAACAATTTGCAAAA	2898
DB	841	TGGGGTCTCAATTTTGCTACAGGAAGAAAACGTGATTCAGATATACATAAACAATTTGCAAAA	900
QY	2899	CAGTGTGAAGTTCCTTCCATCAGAGGAACTGTGTATATGTACCTTTGGGCTCATAGCTAAA	2958
DB	901	CAGTGTGAAGTTCCTTCCATCAGAGGAACTGTGTATATGTACCTTTGGGCTCATAGCTAAA	960
QY	2959	ACCAAAACAAGGCTGTGATATCTTAAATGTCAAACTGGGATGCTGTGAGGCATAGTCGC	3018
DB	961	ACCAAAACAAGGCTGTGATATCTTAAATGTCAAACTGGGATGCTGTGAGGCATAGTCGC	1020
QY	3019	AAACATCTGTGGCAGTGGTTCACAGATGATGTGTGAAACAACTCTGTAAATGAACCTTTCACT	3078
DB	1021	AAACATCTGTGGCAGTGGTTCACAGATGATGTGTGAAACAACTCTGTAAATGAACCTTTCACT	1080
QY	3079	ATCCAAAGCACTCTAAGTTTGAACCTCGAGTCAACAGAGCTCTAGACATAATAGTGAAGT	3138
DB	1081	ATCCAAAGCACTCTAAGTTTGAACCTCGAGTCAACAGAGCTCTAGACATAATAGTGAAGT	1140
QY	3139	GAATCTGTGCCATCGAGTATGTTTCATATTTGGAGATGACCGGTTTGGCAGAGCTCTACT	3198
DB	1141	GAATCTGTGCCATCGAGTATGTTTCATATTTGGAGATGACCGGTTTGGCAGAGCTCTACT	1200
QY	3199	AGTACATTTTCTTTGATATCAATGAAGATACAGAGCAACATTTTATGACCGATCTGGA	3258
DB	1201	AGTACATTTTCTTTGATATCAATGAAGATACAGAGCAACATTTTATGACCGATCTGGA	1260
QY	3259	CCCATAAAGGATAAAAAATTCATTTCCCTTTCTTTGCTTCTAGTAAACTTTGGAAGATCGT	3318
DB	1261	CCCATAAAGGATAAAAAATTCATTTCCCTTTCTTTGCTTCTAGTAAACTTTGGAAGATCGT	1320
QY	3319	ATCTTAAATTCGCTTACTTTTGCTTAAACAAAAAACATCGTAGTAGCAGTGATCCAAAAGGA	3378
DB	1321	ATCTTAAATTCGCTTACTTTTGCTTAAACAAAAAACATCGTAGTAGCAGTGATCCAAAAGGA	1380
QY	3379	GGGAAATTTATCATCTGAAGTAAACAGCAAGCAAGCGGAATCGAAACACTTTACGGAGCCC	3438
DB	1381	GGGAAATTTATCATCTGAAGTAAACAGCAAGCAAGCGGAATCGAAACACTTTACGGAGCCC	1440
QY	3439	AGTGTGATTTTAAATCATAGTGAATGATTTTACACCCATATCCACTGTACAGAAAAATTA	3498
DB	1441	AGTGTGATTTTAAATCATAGTGAATGATTTTACACCCATATCCACTGTACAGAAAAATTA	1500
QY	3499	CAATTAGAGCTTCATTTATGGGAAATAAGCAATGAAGACACTGTGTAGTACACCAAGC	3558
DB	1501	CAATTAGAGCTTCATTTATGGGAAATAAGCAATGAAGACACTGTGTAGTACACCAAGC	1560
QY	3559	ATTCGAGAAATGACTTAAANATCCCAAGNAATTTTGGTACAGAGATCAAGAGAAAT	3618
DB	1561	ATTCGAGAAATGACTTAAANATCCCAAGNAATTTTGGTACAGAGATCAAGAGAAAT	1620
QY	3619	ACAAGCCGAGAGAGGTTAGTAGTAGAAGTTTCAACAGAGCTCACATATGAAGATACGTAGC	3678
DB	1621	ACAAGCCGAGAGAGGTTAGTAGTAGAAGTTTCAACAGAGCTCACATATGAAGATACGTAGC	1680
QY	3679	CAAAAGTTTCAATACAGACACTCAACAAAGTGGCATAAAGTTCAATGAGCTCAAGTCTTCA	3738
DB	1681	CAAAAGTTTCAATACAGACACTCAACAAAGTGGCATAAAGTTCAATGAGCTCAAGTCTTCA	1740

Qy	3739	CGAGAGACAGTAGGTGTAGATGCTTAACTATATGGAACACAGACTGTGTGAACGATGAGTACT	3799
Db	1741	CGAGAGACAGTAGGTGTAGATGCTTAACTATATGGAACACAGACTGTGTGAACGATGAGTACT	1800
Qy	3799	GTGGTAGTACTAAACTATTTAAGACAAGCCACTATTTTGAGCGCCACAGTCTTAACCATCTG	3858
Db	1801	GTGGTAGTACTAAACTATTTAAGACAAGCCACTATTTTGAGCGCCACAGTCTTAACCATCTG	1860
Qy	3859	TCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCTCCAGGTTCTTCTCATAGGTTTCCT	3918
Db	1861	TCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCTCCAGGTTCTTCTCATAGGTTTCCT	1920
Qy	3919	AGAAGACACAGTCCCTTTAAAGCACCCCTCTATTGTCTCAATTAAGTCTAGCAGATTGT	3978
Db	1921	AGAAGACACAGTCCCTTTAAAGCACCCCTCTATTGTCTCAATTAAGTCTAGCAGATTGT	1980
Qy	3979	AACHTTAGTTACACAAGTTCCTAGAGATGCTTTTGGCTATGCTACACTGAAGAAAGCTACAG	4038
Db	1981	AACHTTAGTTACACAAGTTCCTAGAGATGCTTTTGGCTATGCTACACTGAAGAAAGCTACAG	2040
Qy	4039	CAACAAAGAATGCATCCATCCTTATCTCACTCTGAAGCTTTTGGCATCTCCAGCAAAAAGAT	4098
Db	2041	CAACAAAGAATGCATCCATCCTTATCTCACTCTGAAGCTTTTGGCATCTCCAGCAAAAAGAT	2100
Qy	4099	GTGCTATTTACTATACCATCACCATGAAGGCCAACAGTTTGTAGTCCAGATTAAACACCA	4158
Db	2101	GTGCTATTTACTATACCATCACCATGAAGGCCAACAGTTTGTAGTCCAGATTAAACACCA	2160
Qy	4159	AGCAGGTTTCATGAAGCCCTTAAGTTATGCAATCAATTAGATTAAGAAGATTTTATTGAGTCCCT	4218
Db	2161	AGCAGGTTTCATGAAGCCCTTAAGTTATGCAATCAATTAGATTAAGAAGATTTTATTGAGTCCCT	2220
Qy	4219	ATTAAATCAAAATACCCCTGCAACCATCTTCTCAGTGCCTGCTCCAGTGCCTGCTCCAGTGCCACA	4278
Db	2221	ATTAAATCAAAATACCCCTGCAACCATCTTCTCAGTGCCTGCTCCAGTGCCTGCTCCAGTGCCACA	2280
Qy	4279	TATGGGGGTTTCAGATGATTACATTGGTCTTTGCTCTCCCGGTGGATATAAATGATATATTC	4338
Db	2281	TATGGGGGTTTCAGATGATTACATTGGTCTTTGCTCTCCCGGTGGATATAAATGATATATTC	2340
Qy	4339	CAGGTAAAGATATTTCCCTATTTTTCAGACAAAAAACAACAACAATGATGATCGAGGT	4398
Db	2341	CAGGTAAAGATATTTCCCTATTTTTCAGACAAAAAACAACAACAATGATGATCGAGGT	2400
Qy	4399	GCAAGACATTTGCCCATGATGAGGAGGTCCTTCCATCTGSAACCTGAGGTCCTTGTAATA	4458
Db	2401	GCAAGACATTTGCCCATGATGAGGAGGTCCTTCCATCTGSAACCTGAGGTCCTTGTAATA	2460
Qy	4459	AAATCTTTTTCACTTGTCTACGACAGCAGATGAGTCTTTACGGAAATAATGAATTCATCCAT	4518
Db	2461	AAATCTTTTTCACTTGTCTACGACAGCAGATGAGTCTTTACGGAAATAATGAATTCATCCAT	2520
Qy	4519	TCAGATGCCCTCTCTGTTTTTGAAGATPACAGAGACACTGGACTPACAGGAACATACAGAT	4578
Db	2521	TCAGATGCCCTCTCTGTTTTTGAAGATPACAGAGACACTGGACTPACAGGAACATACAGAT	2580
Qy	4579	GATAACTGCCCTTTATTTGTGTCTGTATTGAAATTTCTGGGTTTTCCAGCCCGACCAACTG	4638
Db	2581	GATAACTGCCCTTTATTTGTGTCTGTATTGAAATTTCTGGGTTTTCCAGCCCGACCAACTG	2640
Qy	4639	AGTCCAATATGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAG	4698
Db	2641	AGTCCAATATGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAG	2700
Qy	4699	ACTATCCATAATCCTTTTGAAGTGGTTCCCTCTAAGTTTTTTCGGGATTTTCTGGATGCAGT	4758
Db	2701	ACTATCCATAATCCTTTTGAAGTGGTTCCCTCTAAGTTTTTTCGGGATTTTCTGGATGCAGT	2760
Qy	4759	GATGGGTGTCTCAAGAAGGCTCAGCTAGCAGCAACCAAGACACAGAAATTTGTTACTAGGT	4818
Db	2761	GATGGGTGTCTCAAGAAGGCTCAGCTAGCAGCAACCAAGACACAGAAATTTGTTACTAGGT	2820
Qy	4819	GTTAAAAACAATTCAGATGATACACCAATGTGCGGTATCTCTCTTCCAAAGAAAGTTCTA	4878

||||| 2821 GTTAAACAAATTCAGATGATACACCAATGCGGTATACCTCTCCAAAGAGTTCTA 2880
Qy 4879 AGATTAGTCATTAAATTTGAGTAGTTCAGTTTCAACTAAATGTCAATGAGACTGGGCTTTTA 4938
Db 2881 AGATTAGTCATTAAATTTGAGTAGTTCAGTTTCAACTAAATGTCAATGAGACTGGGCTTTTA 2940
Qy 4939 ACAATTAAGGAGAGTATCCTCAAAACATTTGATGACATATGCTTAACTCTGAGGTTTCC 4998
Db 2941 ACAATTAAGGAGAGTATCCTCAAAACATTTGATGACATATGCTTAACTCTGAGGTTTCC 3000
Qy 4999 CATTTCGTGCACACTGCACATTCAGACTTCCTCGTTCGGAGTTTCATACAGATTAATTT 5058
Db 3001 CATTTCGTGCACACTGCACATTCAGACTTCCTCGTTCGGAGTTTCATACAGATTAATTT 3060
Qy 5059 CAAGATGTACAGTTTCTACAAATGCATGAAGACAGAGGCTGTGTTGGCAACACCA 5118
Db 3061 CAAGATGTACAGTTTCTACAAATGCATGAAGACAGAGGCTGTGTTGGCAACACCA 3120
Qy 5119 AAGCAACTATAGTTGATACATCTGCTGAATTCCTGAACCTCATATTTATGATGGATATAGA 5178
Db 3121 AAGCGACTATAGTTGATACATCTGCTGAATTCCTGAACCTCATATTTATGATGGATATAGA 3180
Qy 5179 TACATCTATATATATTTGATATTTGATGATTTCTTAAAGGCTCAGAAAATACGACTGAC 5238
Db 3181 TACATCTATATATATTTGATGATTTTCTTAAAGGCTCAGAAAATACGACTGAC 3240
Qy 5239 TAGGCAGCAAGACAGAGGATCTTCTGTACACATGTTCCGAGTTTCTGCTGATGAACA 5298
Db 3241 TAGGCAGCAAGACAGAGGATCTTCTGTACACATGTTCCGAGTTTCTGCTGATGAACA 3300
Qy 5299 GTTGGAACTGCTGACTTTCCTTAAACCAAAACAACTTCTCTCTCTTTGAGCCCTTTT 5358
Db 3301 GTTGGAACTGCTGACTTTCCTTAAACCAAAACAACTTCTCTCTCTTTGAGCCCTTTT 3360
Qy 5359 GAGGGTTCATGATTCATTACCAAGTTTAAAGGTTTCAGTTTACCAATGATGCAAGAG 5418
Db 3361 GAGGGTTCATGATTCATTACCAAGTTTAAAGGTTTCAGTTTACCAATGATGCAAGAG 3420
Qy 5419 CCAAGCACTGAATACCTACATAGTTTCTATTTTCTTCAATTTTAAAGCGTAATGACA 5478
Db 3421 CCAAGCACTGAATACCTACATAGTTTCTATTTTCTTCAATTTTAAAGCGTAATGACA 3480
Qy 5479 GTGGAACAATAATGGGATATGCAAGACACCTTCAAGTTTATTTCTGAATGA- TTTTA 5537
Db 3481 GTGGAACAATAATGGGATATGCAAGACACCTTCAAGTTTATTTCTGAATGATTTTA 3540
Qy 5538 GGTTAAATATACAGATGCTTGTATGTTAACTAACTTTGGAAGCAGGAATCAGTGTC 5597
Db 3541 GGTTAAATATACAGATGCTTGTATGTTAACTAACTTTGGAAGCAGGAATCAGTGTC 3600
Qy 5598 TCTAAGGCTGCATCCTATTACCAATGGGTTGTCTATAAATCTGGCTGTTATTAGAGAG 5657
Db 3601 TCTAAGGCTGCATCCTATTACCAATGGG- TGTGCTATAACT-GCTGGTATTAGAGAG 3658
Qy 5658 GGAAC 5662
Db 3659 GGAAC 3663

RESULT 10
LOCUS CQ411997/c
DEFINITION Sequence 19068 from Patent WO0170979.
ACCESSION CQ411997
VERSION CQ411997.1 GI:41319778
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Lee, J. and Lillie, J.

TITLE Genes, compositions, kits, and method for identification,
JOURNAL assessment, prevention, and therapy of ovarian cancer
FEATURES Patent: WO 0170979-A 19068 27-SEP-2001;
source Millennium Pharmaceuticals, Inc. (US)
1. 4657
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 39.6%; Score 2244.2; DB 6; Length 4657;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2386; Conservative 0; Mismatches 3; Indels 100; Gaps 4;
Qy 3175 GACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCTTCATATCAATGAGATACAGAG 3234
Db 4634 GACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCTTCATATCAATGAGATACAGAG 4575
Qy 3235 CCAACATTTTATGACCGATCTGGACCCATTAAGGATAAAAAATTCATTCCTCTTCTTGCT 3294
Db 4574 CCAACATTTTATGACCGATCTGGACCCATTAAGGATAAAAAATTCATTCCTCTTCTTGCT 4515
Qy 3295 TCTAGTAACTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCCTTAAACAAACAT 3354
Db 4514 TCTAGTAACTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCCTTAAACAAACAT 4455
Qy 3355 CGTAGTAGCAGTCATCCAAAGGAGGGAATTAATCATCTCAAGATAGACAGCAACAGG 3414
Db 4454 CGTAGTAGCAGTCATCCAAAGGAGGGAATTAATCATCTCAAGATAGACAGCAACAGG 4395
Qy 3415 CGAATCAGAACACCTTACGGAGCCAGTGTGTTGATTTTAAATCATATAGTATGATTTTACACCC 3474
Db 4394 CGAATCAGAACACCTTACGGAGCCAGTGTGTTGATTTTAAATCATATAGTATGATTTTACACCC 4335
Qy 3475 ATATCCACTGTA CAGAAAAATTAATTAAGACATTTCAATTTATGGGAATAAGACAT 3534
Db 4334 ATATCCACTGTA CAGAAAAATTAATTAAGACATTTCAATTTATGGGAATAAGACAT 4275
Qy 3535 GRAGACACTGTTAGTACACCAAGCATTTGGGAATTAATTAATTTTCAACCAAGATTTT 3594
Db 4274 GAAGACACTGTTAGTACACCAAGCATTTGGGAATTAATTAATTTTCAACCAAGATTTT 4215
Qy 3595 GGTACAGAGAAATCACAGAGAAAAATCAAGCCGAGAGAGGTTAGTAGTAGAAGTTCAACG 3654
Db 4214 GGTACAGAGAAATCACAGAGAAAAATCAAGCCGAGAGAGGTTAGTAGTAGAAGTTCAACG 4155
Qy 3655 AGCTCATATGAAGATACGTAGCCAAAAGTTTCAATACAGACACTACAA CAAGTGGCA 3714
Db 4154 AGCTCATATGAAGATACGTAGCCAAAAGTTTCAATACAGACACTACAA CAAGTGGCA 4095
Qy 3715 AGTTCAATGAGCTCAAGTCTTCCAGAGAGACAGTAGGTAGTATGCTACAACTATGGAC 3774
Db 4094 AGTTCAATGAGCTCAAGTCTTCCAGAGAGACAGTAGGTAGTATGCTACAACTATGGAC 4035
Qy 3775 ACAGACTGTGGAAGACATGAGTACTGTGTAAGTACTAAAAATTAAGACAGCAACTAT 3834
Db 4034 ACAGACTGTGGAAGACATGAGTACTGTGTAAGTACTAAAAATTAAGACAGCAACTAT 3894
Qy 3835 TTGACGCCACAGTCTAAACCATCTGTCTCTCTCCAAATCAAAATTCGGGTGTCCTCGTGCCT 3894
Db 4008 ----- 4009
Qy 3895 CCAGGTTCTTCTCATAGCTTCTTCTAGAGAGACAGTCCCTTAAAGCACCTCTATTGCT 3954
Db 4008 ---GGTTCTTCTCATAGCTTCTTCTAGAGAGACAGTCCCTTAAAGCACCTCTATTGCT 3952
Qy 3955 ACAATTTAAAGTCTAGCAGATTTAACTTTAGTTTACAGAGTTCTAGAGATGCTTTTGGC 4014
Db 3951 ACANTTAAAGTCTAGCAGATTTAACTTTAGTTTACAGAGTTCTAGAGATGCTTTTGGC 3892
Qy 4015 TATGCTACACTGAAAAAGACTTACAGCAACAAAGAAATGCATCCATCCTTATCTCACTCGAA 4074
|||||

Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
Patent: WO 0170979-A 19068 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. 4657
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Db 3891 TATGCTACACTGAAAGACTACAGCAACAAGAAATGCATCCATCTTATCTCACTCTGAA 3832
QY 4075 GCTTTGGCATCTCCAGCAAAAGATGCTATTTACTGATACCATCACCATGAGGCCAAC 4134
Db 3831 GCTTTGGCATCTCCAGCAAAAGATGCTATTTACTGATACCATCACCATGAGGCCAAC 3772
QY 4135 AGTTTTGAGTCCAGATTAACACCAAGCAGGTTTCATGAAAGCCCTTAAGTTATGTCATCATTA 4194
Db 3771 AGTTTTGAGTCCAGATTAACACCAAGCAGGTTTCATGAAAGCCCTTAAGTTATGTCATCATTA 3712
QY 4195 GATAAAGAAGATTTATGAGTCTCTTAATCAAAATACCCTGCAAGCATCTTCCCTCAGTG 4254
Db 3711 GATAAAGAAGATTTATGAGTCTCTTAATCAAAATACCCTGCAAGCATCTTCCCTCAGTG 3652
QY 4255 CGTCCATCGTCCAGTCCCATATGAGGTTTCAGATGATTAATGTTGGTCTTGTCTCTC 4314
Db 3651 CGTCCATCGTCCAGTCCCATATGAGGTTTCAGATGATTAATGTTGGTCTTGTCTCTC 3592
QY 4315 CCGGTGGATATAAATGATATATTTCCAGGTAAAGGATATTCCTTATTTTCAGACAAAAAAC 4374
Db 3591 CCGGTGGATATAAATGATATATTTCCAGGTAAAGGATATTCCTTATTTTCAGACAAAAAAC 3532
QY 4375 ATACCAACCATGATGATCGAGTCAAGAGCAATTTGCCATGATCAGGAGGTCTTCCA 4434
Db 3531 ATACCAACCATGATGATCGAGTCAAGAGCAATTTGCCATGATCAGGAGGTCTTCCA 3472
QY 4435 TCTGGAACCTGGAGGTCTTCTGTAATAAATCTTTTCACTTCTGCTACGACAGATGAGTCTT 4494
Db 3471 TCTGGAACCTGGAGGTCTTCTGTAATAAATCTTTTCACTTCTGCTACGACAGATGAGTCTT 3412
QY 4495 ACGGAAATTAATGAATTCATTCATTCAGATGCTCTCTGTTTTTAAAGATACAGAAAGC 4554
Db 3411 ACGGAAATTAATGAATTCATTCATTCAGATGCTCTCTGTTTTTAAAGATACAGAAAGC 3352
QY 4555 ACTGGAATACAGGAACATACAGATGATAAATCTGCTTTATGTTCTGTTATGAAATCTG 4614
Db 3351 ACTGGAATACAGGAACATACAGATGATAAATCTGCTTTATGTTCTGTTATGAAATCTG 3292
QY 4615 GGTTCACAGCCAGCAACCACTGAGTGCAATATGATGATTCATTCAGACTTTCAAGATATT 4674
Db 3291 GGTTCACAGCCAGCAACCACTGAGTGCAATATGATGATTCATTCAGACTTTCAAGATATT 3232
QY 4675 CCATATTTCTGATGTTGTTGAGCAGACTATCCATAATCTTTAGAAAGTGGTTCCTCTAAG 4734
Db 3231 CCATATTTCTGATGTTGTTGAGCAGACTATCCATAATCTTTAGAAAGTGGTTCCTCTAAG 3172
QY 4735 TTTTCGGGATTTCTGGATGCAATGAGGTTGCTCAAGAAAGGCTCAGTACGAGCACC 4794
Db 3171 TTTTCGGGATTTCTGGATGCAATGAGGTTGCTCAAGAAAGGCTCAGTACGAGCACC 3112
QY 4795 AAAAGCACAGAATTTGTTACTAGTGTAAACAATTCAGAGTATACCAATGATGCGGT 4854
Db 3111 AAAAGCACAGAATTTGTTACTAGTGTAAACAATTCAGAGTATACCAATGATGCGGT 3052
QY 4855 ATACTCTTTCCGAAAGAAGTTCTAAGATGATGATCAATTAATTTGAGTAGTTTCACT 4914
Db 3051 ATACTCTTTCCGAGAGAAGTTCTAAGATGATGATCAATTAATTTGAGTAGTTTCACT 2992
QY 4915 AAATGTCATGAGATGCGGCTTTTAAACAATTAAGGAAGTATCTTCAAAACATTTGATGAC 4974
Db 2991 AAATGTCATGAGATGCGGCTTTTAAACAATTAAGGAAGTATCTTCAAAACATTTGATGAC 2932
QY 4975 ATATGCTTTACTCTGAGGTTTCCATTTGCTGTCACACTGACATTCAGACTCCGCTG 5034
Db 2931 ATATGCTTTACTCTGAGGTTTCCATTTGCTGTCACACTGACATTCAGACTCCGCTG 2872
QY 5035 CGGAGGTTTCATCAAGAATTTATTTCAAGATGTACAGTTTCTACAAATGATGAAAGCA 5094
Db 2871 CGGAGGTTTCATCAAGAATTTATTTCAAGATGTACAGTTTCTACAAATGATGAAAGCA 2812
QY 5095 GAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTGAATCTCTGA 5154
Db 2811 GAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTGAATCTCTGA 2752

QY 5155 CCTCATATTTATGATGGATATAGATACATATATATATATATATATTTGTGATTTCCCTA 5214
Db 2751 CCTCATATTTATGATGGATATAGATACATATATATATATATATTTGTGATTTCCCTA 2692
QY 5215 AAAGCCTCAGAAAAATACGACTGACTAGGAGCAAAAGACAGAGGATATCTTCTGTCACTGT 5274
Db 2691 AAAGCCTCAGAAAAATACGACTGACTAGGAGCAAAAGACAGAGGATATCTTCTGTCACTGT 2632
QY 5275 TCCGAGTTTACTGTTGATCATGAACAGTTGGAACCTGCTGACTTTTCTTAAACCAAACTTC 5334
Db 2631 TCCGAGTTTACTGTTGATCATGAACAGTTGGAACCTGCTGACTTTTCTTAAACCAAACTTC 2572
QY 5335 CTTCTCTCTTTGTTGAGCCTTTTTCAGGAGTTTCATGATTCATTACACAGTTTAAAGAGT 5394
Db 2571 CTTCTCTCTTTGTTGAGCCTTTTTCAGGAGTTTCATGATTCATTACACAGTTTAAAGAGT 2512
QY 5395 TTCACTTACCATTGATGCAAGAGCCAAAGCACTGAATACCTACATAGGTTTCTATTTTC 5454
Db 2511 TTCACTTACCATTGATGCAAGAGCCAAAGCACTGAATACCTACATAGGTTTCTATTTTC 2452
QY 5455 TTTCAATTTTAAAGCGTAAATGACAGTGAACCAATTAATGGGATATGCAAGCACCTTCA 5514
Db 2451 TTTCAATTTTAAAGCGTAAATGACAGTGAACCAATTAATGGGATATGCAAGCACCTTCA 2392
QY 5515 CAAAGTTATTTCTGAATGA-TTTTAGGGTAAATAATACAGATGCTTGTATGTTAACTAAC 5573
Db 2391 CAAAGTTATTTCTGAATGAATTTTAGGGTAAATAATACAGATGCTTGTATGTTAACTAAC 2332
QY 5574 TTGTGAAAGCAGGAAATCAGTGTCTTAAAGGCTGATCCTATTACCAATGGGTTGTG 5633
Db 2331 TTGTGAAAGCAGGAAATCAGTGTCTTAAAGGCTGATCCTATTACCAATGGG-TGTG 2273
QY 5634 CTATTAACCTGGCTGTTATGAGAGGGAAC 5662
Db 2272 CTATAACT-GCTGGTATTAGAGAGGAAC 2245

RESULT 11
CQ488314/c
LOCUS CQ488314 4601 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 20181 from Patent WO0160860.
ACCESSION CQ488314
VERSION CQ488314.1 GI:41453933
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Schlegel R., Endege W.O. and Monahan J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 20181 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..4601
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 39.6%; Score 2239.4; DB 6; Length 4601;
Best Local Similarity 95.7%; Fred. No. 0;
Matches 2383; Conservative 0; Mismatches 6; Indels 100; Gaps 4;
QY 3175 GACCGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTGTGATATCAATGAAGATACAG 3234
Db 4580 GACCGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTGTGATATCAATGAAGATACAG 4521
QY 3235 CCAACATTTTATGACCGATCTGGACCCCATAAAGGATAAAATTCATTCCTTTCTTTGCT 3294
Db 4520 CCAACATTTTATGACCGATCTGGACCCCATAAAGGATAAAATTCATTCCTTTCTTTGCT 4461

QY 3295 TCTAGTAAACCTTGTAAGAATCGTATCTTTAAATTCGCTTACTTTGCTCCTTAACAAAAACAT 3354
DB 4460 TCTAGTAAACCTTGTAAGAATCGTATCTTTAAATTCGCTTACTTTGCTCCTTAACAAAAACAT 4401
QY 3355 CGTAGTAGCAGTGATCCAAAGAGGAGGAAATTAATCATCTGAAAGTAAGACAAAGCAGG 3414
DB 4400 CGTAGTAGCAGTGATCCAAAGAGGAGGAAATTAATCATCTGAAAGTAAGACAAAGCAGG 4341
QY 3415 CGAATCAGAACACATTACCGAGCCAGGTGTTGATTTTAAATCATAGTGTGATGATTTTACACCC 3474
DB 4340 CGAATCAGAACACATTACCGAGCCAGGTGTTGATTTTAAATCATAGTGTGATGATTTTACACCC 4281
QY 3475 ATATCCACTGTACAGAAAAACATTAACAATTAGACTTCAATTTATGAGAAATTAAGACATTT 3534
DB 4280 ATATCCACTGTACAGAAAAACATTAACAATTAGACTTCAATTTATGAGAAATTAAGACATTT 4221
QY 3535 GAGACACTGGTAGTACCAAGCATTTGGAGAAATGACTTAAATTTACCAAGAAATTTT 3594
DB 4220 GAGACACTGGTAGTACCAAGCATTTGGAGAAATGACTTAAATTTACCAAGAAATTTT 4161
QY 3595 GGTAACAGAAATCACAGAGAAAATACAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACG 3654
DB 4160 GGTAACAGAAATCACAGAGAAAATACAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACG 4101
QY 3655 AGCTCAATATGAAGATACGTAGCCAAAGTTTCAATACAGACTACAAACAAGTGGCAAT 3714
DB 4100 AGCTCAATATGAAGATACGTAGCCAAAGTTTCAATACAGACTACAAACAAGTGGCAAT 4041
QY 3715 AGTTCAATAGCTCAAGTCTTCCAGAGAGACAGTAGGTGTAGTAGTACAACTATGGAC 3774
DB 4040 AGTTCAATAGCTCAAGTCTTCCAGAGAGACAGTAGGTGTAGTAGTACAACTATGGAC 3981
QY 3775 ACAGACTGTGGAAGCAGTAGTGTGTAAGTACTAAACTATTAAAGACAAAGCCACTAT 3834
DB 3980 ACAGACTGTGGAAGCAGTAGTGTGTAAGTACTAAACTATTAAAGACAAAGCCACTAT 3955
QY 3835 TTGACGCCACAGCTTAACCAATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCT 3894
DB 3954 ----- 3955
QY 3895 CCAGGTTCTTCTCATACGCTTCTAGNAGAGCAGTCCCTTAAAGCACCTCTATGTCT 3954
DB 3954 ---GGTTCCTTCGTCAGGCTTCTAGNAGAGCAGTCCCTTAAAGCACCTCTATGTCT 3898
QY 3955 ACAATTAAGAGTCTAGCAGATTTGAATTTAGTTTACAAAGTCTAGAGATGCTTTTGGC 4014
DB 3897 ACAATTAAGAGTCTAGCAGATTTGAATTTAGTTTACAAAGTCTAGAGATGCTTTTGGC 3838
QY 4015 TATGCTACACTGAAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTCTGAA 4074
DB 3837 TATGCTACACTGAAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTCTGAA 3778
QY 4075 GCTTTGGCATCTCAGCAAAAGATGTCTATTTAGTATACATCAACATGAAAGGCCAAC 4134
DB 3777 GCTTTGGCATCTCAGCAAAAGATGTCTATTTAGTATACATCAACATGAAAGGCCAAC 3718
QY 4135 AGTTTGGTCCAGATTAAACCAAGCAGGTTCATGAAAGCCTTAAAGTTATGATCATTTA 4194
DB 3717 AGTTTGGTCCAGATTAAACCAAGCAGGTTCATGAAAGCCTTAAAGTTATGATCATTTA 3658
QY 4195 GATAAAGAGATTTATTTAGTCTCTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTG 4254
DB 3657 GATAAAGAGATTTATTTAGTCTCTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTG 3598
QY 4255 CGGTCAATGTGTCCAGTGCCACATATGGGGGTTCAGATGATTTACATTTGCTGTCTC 4314
DB 3597 CGGTCAATGTGTCCAGTGCCACATATGGGGGTTCAGATGATTTACATTTGCTGTCTC 3538
QY 4315 CCGGTGATATAATGATATTTCCAGGTAAGGATATTTCCCTATTTTTCAGACAAAAC 4374
DB 3537 CCGGTGATATAATGATATTTCCAGGTAAGGATATTTCCCTATTTTTCAGACAAAAC 3478

QY 4375 ATACCAACCAATGATGATCGAGGTGCAAGACATTTGCCCATGATGCAAGAGGCTTTCCA 4434
DB 3477 ATACCAACCAATGATGATCGAGGTGCAAGACATTTGCCCATGATGCAAGAGGCTTTCCA 3418
QY 4435 TCTGGAACTGGAGGTCTTGTAAAAATTTCTTTTCACTTGTCTACGACAGCAGATGATCTT 4494
DB 3417 TCTGGAACTGGAGGTCTTGTAAAAATTTCTTTTCACTTGTCTACGACAGCAGATGATCTT 3358
QY 4495 ACCGAATAATGAATTCATTCATTCAGATGCCTCTCTGTTTGTAGAAAGTACAGAAGAC 4554
DB 3357 ACCGAATAATGAATTCATTCATTCAGATGCCTCTCTGTTTGTAGAAAGTACAGAAGAC 3298
QY 4555 ACTGGACTACAGAAACATACAGATGAATACTGCTTTTATTTGTCTGTATTTGAAATTTCTG 4614
DB 3297 ACTGGACTACAGAAACATACAGATGAATACTGCTTTTATTTGTCTGTATTTGAAATTTCTG 3238
QY 4615 GGTTCGAGCCCAAGCACTGAGTGCAATATGATGATTCATTCAGACTTTTCAAGATATT 4674
DB 3237 GGTTCGAGCCCAAGCACTGAGTGCAATATGATGATTCATTCAGACTTTTCAAGATATT 3178
QY 4675 CCATATTTCTGATTTGGTGTGAGCAGACTATCCATAATCTTTTGAAGTGGTTCCCTCTAAG 4734
DB 3177 CCATATTTCTGATTTGGTGTGAGCAGACTATCCATAATCTTTTGAAGTGGTTCCCTCTAAG 3118
QY 4735 TTTTCGGGGATTTCTGGATGCACTGGGTGTCTCAAGAAGCTCAGCTAGCAGACCC 4794
DB 3117 TTTTCGGGGATTTCTGGATGCACTGGGTGTCTCAAGAAGCTCAGCTAGCAGACCC 3058
QY 4795 AAAAGCACAGAAATTTGTTACTAGGTGTTAAACAATTTCCAGATGATACCAATGTGCGCT 4854
DB 3057 AAAAGCACAGAAATTTGTTACTAGGTGTTAAACAATTTCCAGATGATACCAATGTGCGCT 2998
QY 4855 ATACTCTCTCGCAAGAAAGTTCTAAGATTAGTCAATTAATTTAGTAGTCTCAGTTTCAACT 4914
DB 2997 ATACTCTCTCGCAAGAAAGTTCTAAGATTAGTCAATTAATTTAGTAGTCTCAGTTTCAACT 2938
QY 4915 AAATGTCATGAGACTGGCTTTTAAACAATTAAGGAGAAGTATCCTCAAAATTTGATGAC 4974
DB 2937 AAATGTCATGAGACTGGCTTTTAAACAATTAAGGAGAAGTATCCTCAAAATTTGATGAC 2878
QY 4975 ATATGCTTTACTCTCAGGTTTCCCATTTGCTGTCACTGCACTTCAGACTTCCGTGT 5034
DB 2877 ATATGCTTTACTCTCAGGTTTCCCATTTGCTGTCACTGCACTTCAGACTTCCGTGT 2818
QY 5035 CGAGGTTTCATACAGAAATTTATTTCAAGATGTACAGTTTCTCAAAATGCAATGAAGAAGCA 5094
DB 2817 CGAGGTTTCATACAGAAATTTATTTCAAGATGTACAGTTTCTCAAAATGCAATGAAGAAGCA 2758
QY 5095 GAGGCTGTGTGGCAACACCAACCAAGCAACCTATATAGTTGATACATCTGCTGAATCCTGA 5154
DB 2757 GAGGCTGTGTGGCAACACCAACCAAGCAACCTATATAGTTGATACATCTGCTGAATCCTGA 2698
QY 5155 CTTATATTTATGATGATATAGATACATATATATATATATATATTTTGTGGATTTTCTTA 5214
DB 2697 CTTATATTTATGATGATATAGATACATATATATATATATATTTTGTGGATTTTCTTA 2638
QY 5215 AAAGCTCAGAAAATACGACTGACTAGGCAAGCAAGAGGAGTCTTCTCTGACACTGT 5274
DB 2637 AAAGCTCAGAAAATACGACTGACTAGGCAAGCAAGAGGAGTCTTCTCTGACACTGT 2578
QY 5275 TCCGAGTTTACTGGTATCATGAACAGTTGGAACCTGCTGACTTTTCTTAAACCAAAACACTTC 5334
DB 2577 TCCGAGTTTACTGGTATCATGAACAGTTGGAACCTGCTGACTTTTCTTAAACCAAAACACTTC 2518
QY 5335 CTTCTCTCTTTGTTGAGGCTTTTGAAGGTTTCAATGATTCATTTACCAAGTTTTAAAGT 5394
DB 2517 CTTCTCTCTTTGTTGAGGCTTTTGAAGGTTTCAATGATTCATTTACCAAGTTTTAAAGT 2458
QY 5395 TTCAGTTTACCATTTGATGCAAGGCAAGCAGTGAATACCTACATAGTGTCTTATTTTC 5454
DB 2457 TTCAGTTTACCATTTGATGCAAGGCAAGCAGTGAATACCTACATAGTGTCTTATTTTC 2398
QY 5455 TTTTCAATTTTAAAGCGTAAATGACAGTGGAAACATAATATGGGATATGCAAGAGCCCTTCA 5514


```
Db 2397 TTTTCATTAAAGCGTAATGACAGTGGACCAATATATGGATATGCGAAGCACCCCTTCA 2338
QY 5515 CAAAGTTATTTCTGAATGA-TTTTAGGGTAAATAATACAGATCCCTGTATGTAACTAAC 5573
Db 2337 CAAAGTTATTTCTGAATGATTTTATAGGGTAAATAATACAGATCCCTGTATGTAACTAAC 2278
QY 5574 TTGTGGAAGCAGGGAATCAGTCTCTCCTAGGCTGCATCCTATTACCAATGGGGTTGTG 5633
Db 2277 TTGTGGAAGCAGGGAATCAGTCTCTCCTAGGCTGCATCCTATTACCAATGGGG-TGTG 2219
QY 5634 CTATAACTGCGTGTATTAGAGAGGGAAC 5662
Db 2218 CTATAACT-GCTGTATTAGAGAGGGAAC 2191

RESULT 12
LOCUS CQ494141/1 4602 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 26008 from Patent WO0160860.
ACCESSION CQ494141
VERSION CQ494141.1 GI:41459760
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 26008 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source 1..4602
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 39.6%; Score 2239.4; DB 6; Length 4602;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 2383; Conservative 0; Mismatches 6; Indels 100; Gaps 4;
QY 3175 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCCCTTGATATCAATGAAGATACAGAG 3234
Db 4580 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCCCTTGATATCAATGAAGATACAGAG 4521
QY 3235 CCAACATTTTATGACCGATCTGGACCCATAAAGGATAAAAATTCATTCCTTTCTTTGCT 3294
Db 4520 CCAACATTTTATGACCGATCTGGACCCATAAAGGATAAAAATTCATTCCTTTCTTTGCT 4461
QY 3295 TCTAGTAAACTTTGGAAGATCGTATCTTAAATTCGCTTTACTTTGCTTAAACAAAAAAT 3354
Db 4460 TCTAGTAAACTTTGGAAGATCGTATCTTAAATTCGCTTTACTTTGCCCTAAACAAAAAAT 4401
QY 3355 CGTAGTACAGTGTATCCAAAGCAGGGAATTTATCATCTGAAGTAAAGCAAGCAACAGG 3414
Db 4400 CGTAGTACAGTGTATCCAAAGCAGGGAATTTATCATCTGAAGTAAAGCAAGCAACAGG 4341
QY 3415 CCAATCAGAACTTTACGGAGCCAGTGTGTGATTTTAAATCATAGTATGATTTTACACCC 3474
Db 4340 CCAATCAGAACTTTACGGAGCCAGTGTGTGATTTTAAATCATAGTATGATTTTACACCC 4281
QY 3475 ATATCCACTGTACAGAAAACATTACAAATTAGAGACTTCAATTTATGGGGAATTAAGCACATT 3534
Db 4281 ATATCCACTGTACAGAAAACATTACAAATTAGAGACTTCAATTTATGGGGAATTAAGCACATT 4221
QY 3535 GAAGACACTGGTAGTACACCAAGCAATTGGAGAAAATGACTTAAATTCACCAAGAAATTTT 3594
Db 4220 GAAGACACTGGTAGTACACCAAGCAATTGGAGAAAATGACTTAAATTCACCAAGAAATTTT 4161
QY 3595 GGTACAGAGAATCACAGAGAAAATCAAGCCGAGAGAGGTTAGTAGTAAAGTTTCAAG 3654
```

```
Db 4160 GGTACAGAGAATCACAGAGAAAATCAAGCCGAGAGAGGTTAGTAGTAAAGTTTCAAGC 4101
QY 3655 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATAAGACACTTACAAAGTGGCATATA 3714
Db 4100 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATAAGACACTTACAAAGTGGCATATA 4041
QY 3715 AGTTCAATGAGCTCAAGTCCCTTACGAGAGACAGTAGGTAGATGCTTACAACTATGAC 3774
Db 4040 AGTTCAATGAGCTCAAGTCCCTTACGAGAGACAGTAGGTAGATGCTTACAACTATGAC 3981
QY 3775 ACAGACTCTGGAAGCATGAGTACTCTGTGTAAGTACTAAAACTATTAAAGCAAGCCACTAT 3834
Db 3980 ACAGACTCTGGAAGCATGAGTACTGT----- 3955
QY 3835 TTGACGCCACAGCTTAACCATCTCTCTCTCCAAATCAAAATTCGGTGTCCCTGTGTGCT 3894
Db 3954 ----- 3955
QY 3895 CCAGGTTCTTCTCATACGCTTCCCTAGAGAGACAGTCCCTTAAAGACCCCTCTATTGCT 3954
Db 3954 ---GGTTCTTCCGTCAGGCTTCCCTAGAGAGACAGTCCCTTAAAGACCCCTCTATTGCT 3898
QY 3955 ACAATTTAAAGTCTAGCAGATTGTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGC 4014
Db 3897 ACAATTTAAAGTCTAGCAGATTGTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGC 3838
QY 4015 TATGCTACACTGAAAGAAGTACAGCAACAAGAAATGCATCCATCTTATCTCACTCTGAA 4074
Db 3837 TATGCTACACTGAAAGAAGTACAGCAACAAGAAATGCATCCATCTTATCTCACTCTGAA 3778
QY 4075 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTCTGATACCATCACCATGAAGGCCAAC 4134
Db 3777 GCTTTGGCATCTCCAGCAAAAGATGTGCTATTCTGATACCATCACCATGAAGGCCAAC 3718
QY 4135 AGTTTGTAGTCCAGATTAAACCAAGCAGGTTTCATGAAGGCTTAAAGTTATGATCATTA 4194
Db 3717 AGTTTGTAGTCCAGATTAAACCAAGCAGGTTTCATGAAGGCTTAAAGTTATGATCATTA 3658
QY 4195 GATAAAGAAGATTTATTGAGTCTTAAATCAAAATACCTGCAACGATCTTCCCTCAGTG 4254
Db 3657 GATAAAGAAGATTTATTGAGTCTTAAATCAAAATACCTGCAACGATCTTCCCTCAGTG 3598
QY 4255 CGGTCCATGTTGTCAGTGCCACATATCGGGGTTTCAGATGATTTACATTTGGTCTTCTCTC 4314
Db 3597 CGGTCCATGTTGTCAGTGCCACATATCGGGGTTTCAGATGATTTACATTTGGTCTTCTCTC 3538
QY 4315 CCGTGGATATAAATGATATATTTCCAGGTAAAGGATATTCCTCTATTTTTCAGACAAAAAC 4374
Db 3537 CCGTGGATATAAATGATATATTTCCAGGTAAAGGATATTCCTCTATTTTTCAGACAAAAAC 3478
QY 4375 ATACCACATGATGATCGAGGTGCAAGAGATTGTCCTATGATGAGAGGCTCTTCCA 4434
Db 3477 ATACCACATGATGATCGAGGTGCAAGAGATTGTCCTATGATGAGAGGCTCTTCCA 3418
QY 4435 TCTGGAATCGGAGGCTTGTAAATAATCTTTTCACTTTGCTACGACAGCAGATGAGTCTT 4494
Db 3417 TCTGGAATCGGAGGCTTGTAAATAATCTTTTCACTTTGCTACGACAGCAGATGAGTCTT 3358
QY 4495 ACAGAAAATAAATGAAATTCATTCATTCAGATGCTCTCTGTTTTTAAAGATACAGAAAC 4554
Db 3357 ACAGAAAATAAATGAAATTCATTCATTCAGATGCTCTCTGTTTTTAAAGATACAGAAAC 3298
QY 4555 ACTGACTACAGGAACATACAGATGATACTGCCCTTTTATTGTTGCTGTATTGAAATTTCTG 4614
Db 3297 ACTGACTACAGGAACATACAGATGATACTGCCCTTTTATTGTTGCTGTATTGAAATTTCTG 3238
QY 4615 GGTTCACAGCCAGCAACCACTGAGTCAATATCTAGTCAATTCAGACTTTTCAAGATATT 4674
Db 3237 GGTTCACAGCCAGCAACCACTGAGTCAATATCTAGTCAATTCAGACTTTTCAAGATATT 3178
QY 4675 CCATATTTCTGATTTGGTGTGAGCAGATATCCATATCTCTTAGAAGTGTTCCTCTCAAG 4734
```

Db	3177	CCATATTCGTGTTGAGCAGACTATCCATAATCTTTTGAAGTGGTCCCTCTTAAG	3118
Qy	4735	TTTTCGGGATTTCTGGATGTCAGTGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCACC	4794
Db	3117	TTTTCGGGATTTCTGGATGTCAGTGGGGTGTCTCAAGAAGGCTCAGCTAGCAGCACC	3058
Qy	4795	AAAGCAGAGAATTTGTTACTAGTGTAAACAAATTCAGATGATACACCAATGTCGGT	4854
Db	3057	AAAGCAGAGAATTTGTTACTAGTGTAAACAAATTCAGATGATACACCAATGTCGGT	2998
Qy	4855	ATACTCCTTCGAAAGAGTTCTAGATTAGTTCATTAAATTTGAGTGTTCAGTTTCAACT	4914
Db	2997	ATACTCCTTCGAAAGAGTTCTAGATTAGTTCATTAAATTTGAGTGTTCAGTTTCAACT	2938
Qy	4915	AAATGTCATGAGCTGGGCTTTTAAATTAAGGAGAGTATCTCTCAACATTTGATGAC	4974
Db	2937	AAATGTCATGAGCTGGGCTTTTAAATTAAGGAGAGTATCTCTCAACATTTGATGAC	2878
Qy	4975	ATATGCTTTTACTCTGAGGTTTCCCATTTGCTGTGCACATGCAATTCAGACTTCCGGT	5034
Db	2877	ATATGCTTTTACTCTGAGGTTTCCCATTTGCTGTGCACATGCAATTCAGACTTCCGGT	2818
Qy	5035	CGGAGGTTTATCAGAGATTTTCAAGATGTACAGTTTCTCAAAATGCATGAAGAACA	5094
Db	2817	CGGAGGTTTATCAGAGATTTTCAAGATGTACAGTTTCTCAAAATGCATGAAGAACA	2758
Qy	5095	GAGGCTGTGTTGGCAACACCAAGCAACCTAGTGTGATACATCTGCTGAATCTTGA	5154
Db	2757	GAGGCTGTGTTGGCAACACCAAGCAACCTAGTGTGATACATCTGCTGAATCTTGA	2698
Qy	5155	CCTCATATTTATGATGATATAGATACATACATATATATATATATATTTTGGATTTCTTA	5214
Db	2697	CCTCATATTTATGATGATATAGATACATACATATATATATATATTTTGGATTTCTTA	2638
Qy	5215	AAAGCTCAGAAATACAGCTAGCAGCAGCAAGACAGAGTATCTTCTGTACACTGT	5274
Db	2637	AAAGCTCAGAAATACAGCTAGCAGCAGCAAGACAGAGTATCTTCTGTACACTGT	2578
Qy	5275	TCCGAGTTACTGTTACATGAACAGTGTGGAAGTCTGCTGCTTCTTAAACCAAAACACTTC	5334
Db	2577	TCCGAGTTACTGTTACATGAACAGTGTGGAAGTCTGCTGCTTCTTAAACCAAAACACTTC	2518
Qy	5335	CTTCTCTCTTTGTTAGCCCTTTTGGGGTTCATGATTCATTACCAAGTTCCTTAAGAGT	5394
Db	2517	CTTCTCTCTTTGTTAGCCCTTTTGGGGTTCATGATTCATTACCAAGTTCCTTAAGAGT	2458
Qy	5395	TTGAGTTACATTTGATGCAAGAGCAAGCACTGAATACCTACATAGATTTTCTATTTTC	5454
Db	2457	TTGAGTTACATTTGATGCAAGAGCAAGCACTGAATACCTACATAGATTTTCTATTTTC	2398
Qy	5455	TTTCATTTTAAAGCGTAATGACAGTGAACAATAATGCGGATATGCAAGCACCCTTCA	5514
Db	2397	TTTCATTTTAAAGCGTAATGACAGTGAACAATAATGCGGATATGCAAGCACCCTTCA	2338
Qy	5515	CAAGTTATTTCTGAATGA- TTTTGAAGGTAATAATACAGATGCCTTGTATGTTAACTAAC	5573
Db	2337	CAAGTTATTTCTGAATGA- TTTTGAAGGTAATAATACAGATGCCTTGTATGTTAACTAAC	2278
Qy	5574	TTGTGGAAGCAGGAATCAGTGTCTTAAGGCTGCATCTTATTACCAATGCGGTTGTG	5633
Db	2277	TTGTGGAAGCAGGAATCAGTGTCTTAAGGCTGCATCTTATTACCAATGCGG- TGTG	2219
Qy	- 5634	CTATACTGGCTGTTATAGAGGGAAC	5662
Db	2218	CTATAACT-GCTGGTATTAGAGGGAAC	2191

RESULT 13
AKI73326
LOCUS AKI73326 6816 bp mRNA linear ROD 28-JUL-2004
DEFINITION Mus musculus mRNA for mKIAA1999 protein.
ACCESSION AKI73326
VERSION AKI73326.1 GI:50511236

KEYWORDS	FLI_CDNA.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	1 Okazaki, N., Kikuno, R.F., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Suga, Y., Seino, S., Nishimura, M., Kaisho, T., Hoshino, K., Kitamura, H., Nagase, T., Ohara, O. and Koga, H. Prediction of the Coding Nucleotide Sequences of Mouse Homologues of KIAA Gene: IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences of cDNA Clones Randomly Sampled from Size-Fractionated Libraries DNA Res. 11, 205-218 (2004) 2 (bases 1 to 6816) Okazaki, N., Kikuno, R.F., Nagase, T., Ohara, O. and Koga, H. Direct Submission Submitted (19-MAY-2004) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918) The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing. FEATURES Location/Qualifiers 1..6816 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="mtg03906" /tissue_type="adult thymus" /note="vector:modified pBC SK+" 1..6816 /gene="mKIAA1999" /notes="CDS is predicted by in silico analysis. Start codon is not identified." /evidence="not_experimental" /product="mKIAA1999 protein" /protein_id="BAD32604.1" /db_xref="GI:50511237" /translation="KLEKWKHKVNSKYVDLIEQLNEALTYRKPIGDGNYVRSNQ RLGRPHVLYGLVHHKTCGLLEVOISITELCHNVRTPLDLDKWDIKKURSL WALNGTSSNWGLNLEQENVIPIILKARQCEVLISIRGTCVYVLIATIKQCCDL KCHSWDSVRHSRKLMPVDPDVEQLCNELSVPTLSLNSSESTSRHSESEAPSS MFLEDDRFGSTSTSTFFLDINEDAPAFYDPGPIKDKNSPFPFGSKLVKNRILNS LTLPTKHRSSDPKGGKLSSENKSNRRIRLTLEPSVDLNHSEDFTSSSAQKSLQLE PSFVGNKHLEDAGSTPSIGENDLKFPKSGFGETHRENTSRERLVEGASSHIKRSQ SFNTDTTSGISSMSSPSPRETVAVDPTAMDTCGLSTVVTXTKVTISHLYTPQSNH LSLUSKNSVSLVPPGSHSLPRRAQSLKAPSIATIKSLADCNFSYTSRDAFGYATLK RLQQRMHPSLSHEALAPKQVLFDTDTTMKANSFESRLTSPRFMKALSASLDE DLSPINHTLQRSSVRSVMSVATYGGSDYIGLALPVDINDI FQIKDVPYFOSKHV PPDDRGARMFSDHAGSLGAGLVKNSPHILRQOMSLTEINNSVHSDASLPLESTE DTGLQHTDNCILYCVCIELGLQPSNQLSSICSHSDLODIPYSDMCEQTHNPLEV PSKPSGSGSGDASQEEKSGSTKTELLGLVKTIPDDTPMCKILLURKEVRLVLNL SSSVTKCHTGLTTIKERYPTQFDICLYSEVSHLSHCTFRLQCRRTFQELFQDQV FLQWBEAEAVLAIPPIQPIVDESAES"
gene	1..6816
CDS	1..6816
Query Match	39.2%; Score 2220.8; DB 10; Length 6816;
Best local Similarity	84.4%; Pred. No. 0;
Matches 2647; Conservative	0; Mismatches 462; Indels 29; Gaps 12;
Qy	2501 CAAAACAATTTGGAAAGTGGCAGGGAATACAACTCCAAATATGTTGACTTGTATTGAGG 2560
Db	1 CAAAACAATTTGGAAAGTGGCAGGGAATATTAATTCAAAGTATGTTGACTTGTATTGAGG 60
Qy	2561 AACAACTCAATGAAGCACTTACTACTTACCGGAAGCCTGTTGATGGTGATAACTATGTTTC 2620
Db	61 AACAACTTAATGAAGCCCTCACTACATACAGGAAGCCTATTTGATGGTGATAACTACGTTTC 120

QY 2621 GTCGGAGTAACCAAGATTACAGCGTCTCACGTCTACCTGCTTATACACCTTTTATGGAC 2680
DB |||||
121 GTCGCGAGTAATCAAGACTACAGCGACCTCATGCTCTACCTGCCGTACACCTTTATGGAC 180
QY 2681 AACTAGTACACCAATAAACHAGGCTGCCATTTGTTGGAAAGTACAGAAATATTTATACAGAAC 2740
DB |||||
181 AACTGGGTACACCAAAACAGGCTGCCATTTTACTGGAAGTACAGAGTATCATCACAGAGC 240
QY 2741 TCTGTGTAATGTTCTGTACACAGATTTCGATAAGTGGGAAGAAATTAAGAACTCAAG 2800
DB |||||
241 TCTGTACAAATGTTGCGACTTCCAGACTTGGACAATATGGGAAGACATTAAGAAAGCTCAAG 300
QY 2801 CATCTCTTTGGGCTTTGGGAAATATCGGCTCATCAAAATTTGGGCTCTCAATTTGCTACAGG 2860
DB |||||
301 CCTCTCTTTGGGCTTTGGGAAATATTTGGCTCATCGAATTTGGGGCTTAATTTGCTACAGG 360
QY 2861 AGAAACGTAATTCAGATATATACTAAACTTTGCAAAACAGTGTGAAGTCTTTTCCATCA 2920
DB |||||
361 AGAAATATGTATCCAGATATATACTAAACTTTGCAAAACAGTGTGAAGTCTTTTCCATCC 420
QY 2921 GAGGACCTGTGTATGTATGTTGCTCTATAGCTTAAGCCAAACAGGCTGTGATATTC 2980
DB |||||
421 GAGGACCTGTGTATGTATGTTGCTCTATAGCTTAAGCCAAACAGGCTGTGATATTT 480
QY 2981 TAAATGTACAACTGGGATGCTGTGAGGCATAGTCGCAAAACATCTGTGSCCAAGTGGTTC 3040
DB |||||
481 TGAATGTACAGCTGGATCTGTGAAGGCAAGTGTGCAAGCATCTATGGCCAGTGGTTC 540
QY 3041 CAGATGATGTGGAAACAACTCTGTAATGAATTTTCATCTATPCCCAAGCACTCTAAGTTGA 3100
DB |||||
541 CAGATGATGTGGAGCAGCTCTGCAATGAGCTCTCATCTGTCTCCCAAGCACTCTCAGCTTGA 600
QY 3101 ACTCGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTGTGCGCATCGAGTATCT 3160
DB |||||
601 ACTCTGAGTGCACAGCTCTAGACATAATAGTGAAGTGAATCTGTGCGCATCGAGTATCT 660
QY 3161 TCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCTTTGATATCA 3220
DB |||||
661 TCATGCTGGAGGACGACCGCTTCGGCAGCAGCTCCACGACCATTTTCTTCTTGATCA 720
QY 3221 ATGAAGATCAGAGCAACAACTTTTATGACCGATCTGGAACCAATAAGGATAAAATTCAT 3280
DB |||||
721 ATGAAGATGTCAGCCAGCATTTTATGACCGACCTGGACCAATAAAGATAAAATTCAT 780
QY 3281 TCCCTTTCTTGTCTCTAGTAACTTGTGAAGTCTGTATCTTAATTTGCTTACTTTGC 3340
DB |||||
781 TCCCTTTTGTGCTCCAGCAAACTTGTAAAGAAATCGTATCTTTAAATTCATTTACTTTGC 840
QY 3341 CTAAACAAAAACATCTGTAGTACAGTGTATCCAAAGAGGAAATTTATCATCTGAAAGTA 3400
DB |||||
841 CTACTAAAAACACCGCAGTGTAGTGTATCCAAAGAGGAAATTCGTCTATCTGAAATA 900
QY 3401 AGCAAGCAACAGCGAAATCAGAACTTACGGAGCCCGAGTGTGATTTTAATCATPATGTG 3460
DB |||||
901 AGCAAGCAACAGCGAAATCAGAACTTACGGAGCCCGAGTGTGATTTGAAATCAGAGC 960
QY 3461 ATGATTTTACACCCATATCCACTGTACAGAAACATTAATTTAGAGACTTCAATTTATGG 3520
DB |||||
961 AGACTTCA---AGCTCCTCAGCAGAAATCATTTACAGCTAGAGCTTCTCTCGTTG 1017
QY 3521 GGAATAGCAGCATTTGAAGCACCTGTGTAGTACACCAAGCAATTTGAGAAAAATGACTTAAAT 3580
DB |||||
1018 GGAATAGCAGCATTTGAAGCACCGGGCAGCACTCCAAGTATTGAGAGAAATGACCTCAAT 1077
QY 3581 TCACCAAGAAATTTGTGTACAGAAATCAGAGAAATCAAGCCGAGAGAGGTTAGTAG 3640
DB |||||
1078 TCCCAAGAGTTTGTGTACCGAGACTCAGAGAAACCAAGCCGAGAAAGGTTGGTTG 1137
QY 3641 TAGAAAGTTCAAGAGCTCAGATATGAAATAGTAGCCAAAGTTTCAATACAGACATA 3700
DB |||||
1138 TGAAGGCTCAGCAAGCTCCCATCAAGATCCGAGCCAGAGCTTTTAAACGGAACATA 1197

QY 3701 CAAACAGTGGCATAAGTTCATAGCTCAAGTCTCTTACGAGAGACAGTAGGTGTAGATG 3760
DB |||||
1198 CGACGAGTGGCATCAGTCCATGAGCTCGAGCCCTCCCGAGAGACAGTCGCTGTGACC 1257
QY 3761 CTACAACTATGACACAGACTGTGGAAGCATCAGTACTCTGTGTAAGTACTATAAACTATTA 3820
DB |||||
1258 CCACAGCTATGACACGGACTGTGGAAGTCTGAGTACCGTGGTAAGTACTATAAACTGTTA 1317
QY 3821 AGCAAGCCACTATTTGAGCCACAGTCTAAACATCTGTCTCTCCAATCAAAATTCG 3880
DB |||||
1318 AGACGAGCCACTATTTGACACACAGTCTAACATCTGTCTCTCCAATCGNACTCAG 1377
QY 3881 TGTCCCTGTGCTCCAGGTTCTTCTCATACGCTTCTTGAAGAGACAGTCCCTTTAAAG 3940
DB |||||
1378 TGTCCCTGTGCTCCAGGTTCTTCTCACACCTTCTTGAAGAGCCAGTCCCTTTAAAG 1437
QY 3941 CACCTCTATTTGCTACAAATTAAGCTTAGCAGATTTGTAACCTTTAGTTACAGAGTTCTA 4000
DB |||||
1438 CGCCCTCCATTTGCAACAAATAAAAGTCTAGCAGATTTGCAACTTCAGCTACACAAAGCTTA 1497
QY 4001 GAGATGCTTTGCTATGCTACACTGAAAGACTACAGCAACAAAGAAATGCAATCCATCT 4060
DB |||||
1498 GAGATGCTTTGCTATGCTACATTTGAAAGGTTACGAGCAAGAAATGCAACCTGCT 1557
QY 4061 TATCTCACTCTGAAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTACTGATACCATCA 4120
DB |||||
1558 TGTCTCACTCGGAAGCTTTGGCTCGCGGCAAAAGATGTCTTTTACCGACACCATCA 1617
QY 4121 CCATGAAGCCAAAGTTTTGTAGTCCAGATTAACACCAAGCAGGTTTCATGAAGCCTTAA 4180
DB |||||
1618 CCATGAAGCCAAAGCTTTGAAATCCAGACTGACGCAAGCAGGTTTATGAAGCTTTAA 1677
QY 4181 GTTATGTCATTTAGATAAAGAAATTTATTCAGTCTCTAATTAATCAAAATACCTCTCAAC 4240
DB |||||
1678 GTTATGCTCATTAAGTAAAGAAATTTACTAAGTCTTAATTAATCAATACCTCCAGC 1737
QY 4241 GATCTTCTCAGTCCGCTCCATGGTGTCCAGTGCACATPATGGGGTTCAGATGATTACA 4300
DB |||||
1738 GCTCTCTCTCAGTGTAGTCTATGGTGTCCAGTGTACATATGGAGGTTCCGATGACTACA 1797
QY 4301 TTGGTCTGCTCTCCGGTGGATATAATGATATATTCCAGGTAAAGGATATCCCTATT 4360
DB |||||
1798 TTGGACTTGTCTCTCCAGTAGACATCAATGACATCTTCCAGATAAAGGATGTTCCTTATT 1857
QY 4361 TTTCAGCAAAAAACATACCAACCATGATGATCGAGGTGCAAGAGCATTTGCCCATGATG 4420
DB |||||
1858 TTCAATCAAGACAGTGCCTCCACAGATGACGAGGTGCAAGATGTTTTCCCATGATG 1917
QY 4421 CAGGAGGCTTCCATCTGGAACCTGGAGGTCCTTTGTAATAAAATTTCTTTTCACTTTGCTAC 4480
DB |||||
1918 GAGCAGGCTTTTCACTGGAGCCGGTGTCTCGTAAAGAACTCTTTTCACTTTGCTCGGC 1977
QY 4481 AGCAGATGATCTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTGTGTTTTAG 4540
DB |||||
1978 AGCAGATGATCTTCCAGAAATAATGAATCTGTCCATTCGAGCGCTCCCTGTTCTTAG 2037
QY 4541 AAAGTACAGAACACATCGGACTACAGGAACATACAGATGATAACTGCCCTTTATTGTGCT 4600
DB |||||
2038 AGAGCACAGAGACATCGGCTGAGGAGCACAGGATGCAATTTGCCCTTTACTGCGTCT 2097
QY 4601 GTATTGAATTTCTGGGTTTTCCAGCCCGCAACCAACTGAGTGCATATGATGATTCATTCAG 4660
DB |||||
2098 GCATTGAATCTCTGGGCTTCCAGCCCGAGTAAACAGCTAAGTTCCATCTCGAGTCACTCCG 2157
QY 4661 ACTTTCAAGATATTCATATTTCTGATTGAGGAGACATATCCATATCTTTTAGAAG 4720
DB |||||
2158 ATCTTTCAAGATATTCGTAATCCGATTTGGTGGAGCAGACGATCCCAACACCTTTAGAA 2217
QY 4721 TGGTTCCTCTAAGTTTTCGGGGATTTCTGGAATGAGTGCATGCGGTGTCTC---AAGAAG 4777
DB |||||
2218 TGGTTCCTCTAAGTTTTCGGGGATTTCTGCGGATTTCTGGCTGAGCGATGCGCATCTCAGGAAG 2277
QY 4778 GCTCAGCTAGCAGCACCAAGACAGAAATTTGTTACTAGGTGTTAAACAAATTCAGATG 4837

QY 4355 CCTATTTTCAGACAAAACATACCACACATGATGATCGAGGTGCAAGAGCATTTGCC 4414
Db 1185 ----- 1184
QY 4415 ATGATCGAGGAGGTCTTCCATCTGGAACGAGGTCTTGTAAAAAANTCTTTCACTTGC 4474
Db 1185 ----- 1184
QY 4475 TACGACAGCAGATGAGTCTTTACGGAATAATGAATTCATTCAGATGCGCTCTCTGT 4534
Db 1185 -----CAGCAGATGAGTCTTTACGGAATAATGAATTCATTCAGATGCGCTCTCTGT 1239
QY 4535 TTTTAGAAAGTACAGAAACACCTGGACTACAGGAACATACAGATGATTAAGTCCCTTTATT 4594
Db 1240 TTTTAGAAAGTACAGAAACACCTGGACTACAGGAACATACAGATGATTAAGTCCCTTTATT 1299
QY 4595 GTGTCGTATTGAAATCTGGGTTTCCAGCCCAACCAACTGAGTGCAATATGTAGTC 4654
Db 1300 GTGTCGTATTGAAATCTGGGTTTCCAGCCCAACCAACTGAGTGCAATATGTAGTC 1359
QY 4655 ATTCAGACTTTCAGATATTCATATTTCTGATGGTGTGAGCAGACTATCCATAATCCTT 4714
Db 1360 ATTCAGACTTTCAGATATTCATATTTCTGATGGTGTGAGCAGACTATCCATAATCCTT 1419
QY 4715 TAGAAGTGGTTCCTCTCTAAAGTTTTCGGGATTTCTGGATGCAATGGGTGTCTCAAG 4774
Db 1420 TAGAAGTGGTTCCTCTCTAAAGTTTTCGGGATTTCTGGATGCAATGGGTGTCTCAAG 1479
QY 4775 AAGGCTCAGCTAGCAGCACCAAAAGCACAGAAATTTTACTAGGTGTTAAAACAATTCAG 4834
Db 1480 AAGGCTCAGCTAGCAGCACCAAAAGCACAGAAATTTTACTAGGTGTTAAAACAATTCAG 1539
QY 4835 ATGATACACCAATGTGCGGTATCTCTTCGCAAGAAGTCTTAGATTAGTCATTAATT 4894
Db 1540 ATGATACACCAATGTGCGGTATCTCTTCGCAAGAAGTCTTAGATTAGTCATTAATT 1599
QY 4895 TGAGTAGTTTCAGTTTCAACTTAATGTCATGAGACTGGGCTTTTAAC 4940
Db 1600 TGAGTAGTTTCAGTTTCAACTTAATGTCATGAGACTGGGCTTTTAAC 1645

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 14:46:37 ; Search time 14461.4 Seconds
(without alignments)
17178.818 Million cell updates/sec

Title: US-10-782-244-2
Perfect score: 5127
Sequence: 1 atggcggcgatcgccgcgg.....atacatctgctgaatctga 5127

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ats.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5127	100.0	5127	6	CQ869396 Sequence
2	5127	100.0	5662	6	CQ869395 Sequence
3	5127	100.0	5662	9	AY515854 Homo sapi
4	4160.4	81.1	9297	10	AY540053 Mus muscu
5	4158.8	81.1	5127	10	AY497009 Mus muscu
6	3829	74.7	3829	6	CQ847976 Sequence
7	3829	74.7	8232	9	AB082530 Homo sapi
8	3800	74.1	8232	6	BD185164 Novel gen
9	3152.8	61.5	5440	9	BC051729 Homo sapi
10	2037.8	39.7	6816	10	AK173326 Mus muscu
11	1992.6	38.9	2430	10	BC022906 Mus muscu
12	1772.8	34.6	4657	6	CQ411997 Sequence
13	1768	34.5	4601	6	CQ488314 Sequence
14	1768	34.5	4602	6	CQ494141 Sequence
15	1494	29.1	1649	6	CQ726625 Sequence
16	1009.4	19.7	97240	9	AC109467 Homo sapi
17	1009.4	19.7	170996	2	AC010938 Homo sapi
18	1009.4	19.7	198748	2	AC016719 Homo sapi
19	874.8	17.1	3129	6	CQ851140 Sequence

20	874.8	17.1	3129	9	AK128358 Homo sapi
21	830.2	16.2	214225	2	AC105581 Rattus no
c 22	763	14.9	153145	10	AC102825 Mus muscu
23	746.6	14.6	1778	9	BC029608 Homo sapi
24	612.8	12.0	765	6	BD020210 Novel gen
25	612.8	12.0	765	6	BD100148 Novel gen
26	534	10.4	534	6	CQ726624 Sequence
27	497.2	9.7	4851	10	BC058643 Mus muscu
28	465.2	9.1	570	6	CQ392940 Sequence
29	465.2	9.1	570	6	CQ399338 Sequence
c 30	385.6	7.5	452	6	CQ477511 Sequence
c 31	383.4	7.5	427	6	CQ498714 Sequence
c 32	383.4	7.5	427	6	CQ507684 Sequence
33	371.6	7.2	973	5	CR391153 Gallus ga
c 34	371	7.2	772	6	BD020211 Novel gen
c 35	371	7.2	772	6	BD100149 Novel gen
36	370	7.2	401	6	AX885079 Sequence
37	370	7.2	401	6	BD024689 Sequence
38	339.8	6.6	4726	6	CQ850194 Sequence
39	339.8	6.6	4726	9	AK127305 Homo sapi
c 40	336.4	6.6	416	6	CQ468343 Sequence
41	334	6.5	341	6	CQ405723 Sequence
c 42	264.6	5.2	164698	2	EX936441 Danio rer
c 43	264.6	5.2	186000	2	CR735111 Danio rer
c 44	196	3.8	188788	2	CR318610 Danio rer
45	194.6	3.8	68372	2	AC023763 Homo sapi

ALIGNMENTS

RESULT 1	CQ869396	Sequence 2 from Patent WO2004074448.	5127 bp	DNA	linear	PAT 13-SEP-2004
LOCUS	CQ869396					
DEFINITION	CQ869396					
ACCESSION	CQ869396					
VERSION	CQ869396.1	GI:51999315				
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	Sabatini, D.M., Kim, D.H., and Sarbassov, D.D.				
AUTHORS		FOR KINASE-ASSOCIATED PROTEINS				
TITLE		Patent: WO 2004074448-A 2 02-SEP-2004;				
JOURNAL		WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)				
FEATURES		Location/Qualifiers				
source		1. 5127				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				

ORIGIN	Query Match	100.0%;	Score 5127;	DB 6;	Length 5127;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 5127;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCGCGGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAAATGAC	60		
Db	1	ATGCGCGGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAAATGAC	60		
Qy	61	AGCGGCGAGAGAACGTCCTGGATCTGACCCGAGAACCTTCTGTAACCTTAAGAGAG	120		
Db	61	AGCGGCGAGAGAACGTCCTGGATCTGACCCGAGAACCTTCTGTAACCTTAAGAGAG	120		
Qy	121	ATTCTCCAAATGTGGCCAGATTCGAGGAGTATCAATATGAGAAAGCTAGGCCATCTG	180		
Db	121	ATTCTCCAAATGTGGCCAGATTCGAGGAGTATCAATATGAGAAAGCTAGGCCATCTG	180		
Qy	181	AATAACTTTACTAAGCTTCTTTGTGATATTGGCCACAGTGAAGAAAACCTGGGCTTTTAC	240		
Db	181	AATAACTTTACTAAGCTTCTTTGTGATATTGGCCACAGTGAAGAAAACCTGGGCTTTTAC	240		

Qy	241	TATAGGATATCATAAATTTGTTGGGTTAGCTTTATTAATGAAGCAAAAGAGTGCGA	300
Db	241	TATAGGATATCATAAATTTGTTGGGTTAGCTTTATTAATGAAGCAAAAGAGTGCGA	300
Qy	301	GCAGAGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGTATTTCTCCAGAAAGGTG	360
Db	301	GCAGAGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGTATTTCTCCAGAAAGGTG	360
Qy	361	CTAAAAATTGAAAGTGGACCTATTATTAATAGCTAGGTGCATTTGATACAAACAGAGCAACGAG	420
Db	361	CTAAAAATTGAAAGTGGACCTATTATTAATAGCTAGGTGCATTTGATACAAACAGAGCAACGAG	420
Qy	421	GTAGAGAGACACAAGCACTTCGATTAGCTCAGAAGATGATTACTGTGGAATGCTTCTCTTG	480
Db	421	GTAGAGAGACACAAGCACTTCGATTAGCTCAGAAGATGATTACTGTGGAATGCTTCTCTTG	480
Qy	481	TTTTCTTAGTCTGTGCACCAACTCATTAAATGTCAGTTGGAAATGATGGACTTCAAGAAAGA	540
Db	481	TTTTCTTAGTCTGTGCACCAACTCATTAAATGTCAGTTGGAAATGATGGACTTCAAGAAAGA	540
Qy	541	GACAGAATGTCGAGCATGCAATGCCATTATCTGTGAACTAGACACTTCAGAATCCAGAG	600
Db	541	GACAGAATGTCGAGCATGCAATGCCATTATCTGTGAACTAGACACTTCAGAATCCAGAG	600
Qy	601	GTGTGSCCTTCGAGGAGGACTTAAACACCATTAATGAAAAAATGTGATTTGCCAATTA	660
Db	601	GTGTGSCCTTCGAGGAGGACTTAAACACCATTAATGAAAAAATGTGATTTGCCAATTA	660
Qy	661	AGTCGAATAAATGAGGCCCTAATTAATCAATTTTGGCACTTCTTAATCATCCAAAGACT	720
Db	661	AGTCGAATAAATGAGGCCCTAATTAATCAATTTTGGCACTTCTTAATCATCCAAAGACT	720
Qy	721	CGGCAGTATGTGCGAGCTGATGTAGAAATPAGAGAGAAATTTAGCACCCCTACTGATTTT	780
Db	721	CGGCAGTATGTGCGAGCTGATGTAGAAATPAGAGAGAAATTTAGCACCCCTACTGATTTT	780
Qy	781	CACATACAGACATAGTCCAGATACAGCTGAGGACAGCTCAAAAGAGACAGAGACGAGA	840
Db	781	CACATACAGACATAGTCCAGATACAGCTGAGGACAGCTCAAAAGAGACAGAGACGAGA	840
Qy	841	TTTCTAGCCAGTAAATGGGAATCATAGCAACATTTCCGATCATGGCGAGGTATTATTAAT	900
Db	841	TTTCTAGCCAGTAAATGGGAATCATAGCAACATTTCCGATCATGGCGAGGTATTATTAAT	900
Qy	901	TTATGTAAACCTGGAATCTGGGATCCAGTCTCTAATAGGAGTACTTTGCATACCRAAT	960
Db	901	TTATGTAAACCTGGAATCTGGGATCCAGTCTCTAATAGGAGTACTTTGCATACCRAAT	960
Qy	961	ATGGAAATAGGCGAGCTACTTGAAGTCTTTATGATATATTTGCTCTCTCTACCT	1020
Db	961	ATGGAAATAGGCGAGCTACTTGAAGTCTTTATGATATATTTGCTCTCTCTACCT	1020
Qy	1021	GTGTGACTGAGGAGTTCATAGAGCACTACTCAGTGTAGATCCAGGGAGGTTCCAGAC	1080
Db	1021	GTGTGACTGAGGAGTTCATAGAGCACTACTCAGTGTAGATCCAGGGAGGTTCCAGAC	1080
Qy	1081	AGTTGGAGCTTTCAGATGGCTTTTGGAGCTGAGGCAAAACATTTCTCTCATCGT	1140
Db	1081	AGTTGGAGCTTTCAGATGGCTTTTGGAGCTGAGGCAAAACATTTCTCTCATCGT	1140
Qy	1141	GCCAGATCCAGGCCAGACCTCATGGAATAATTTTGGCACTGATCTCTGCAATTAAT	1200
Db	1141	GCCAGATCCAGGCCAGACCTCATGGAATAATTTTGGCACTGATCTCTGCAATTAAT	1200
Qy	1201	CGTAATGACATTTTAGAGGCTAGTTGAAGTGAATACAAACAGTCAATCATCTCA	1260
Db	1201	CGTAATGACATTTTAGAGGCTAGTTGAAGTGAATACAAACAGTCAATCATCTCA	1260
Qy	1261	GTTTAGAGCTTACCATCCTTTTAGGAGAGCTTTTATCATATGGCAACAAATTTCTCTCAT	1320
Db	1261	GTTTAGAGCTTACCATCCTTTTAGGAGAGCTTTTATCATATGGCAACAAATTTCTCTCAT	1320
Qy	1321	TCACATAGCCATCAITTTACATGCTTGGCAACCCCTAATGAATATGGCTGCATCTTTGAT	1380
Db	1321	TCACATAGCCATCAITTTACATGCTTGGCAACCCCTAATGAATATGGCTGCATCTTTGAT	1380
Qy	1381	ATCCCCAGAAAGAGACTGCGAGCGCTGAGCTTGAATCTGTTTAAAAACGCTTCCAT	1440
Db	1381	ATCCCCAGAAAGAGACTGCGAGCGCTGAGCTTGAATCTGTTTAAAAACGCTTCCAT	1440
Qy	1441	GAAATGAAGAAAACGAGCACTAAGCCTTATAGTCTTCTCAATTTAGACCACATTTATCGAGAA	1500
Db	1441	GAAATGAAGAAAACGAGCACTAAGCCTTATAGTCTTCTCAATTTAGACCACATTTATCGAGAA	1500
Qy	1501	GCAATTCGACACACACAGAAAACGAGTCAATCTCCGAGTTCAGAAAGATATATTTATC	1560
Db	1501	GCAATTCGACACACACAGAAAACGAGTCAATCTCCGAGTTCAGAAAGATATATTTATC	1560
Qy	1561	CTTAAGGATACAGAGNAAGCTCTTTTAAATTAACCTTTAGAGTAGCAAGTCTTCAACAT	1620
Db	1561	CTTAAGGATACAGAGNAAGCTCTTTTAAATTAACCTTTAGAGTAGCAAGTCTTCAACAT	1620
Qy	1621	AAAGAGAACTTTGAAATGGAATCTTATAGGGAACATTTCTTAAGTGGCCAAATGTA	1680
Db	1621	AAAGAGAACTTTGAAATGGAATCTTATAGGGAACATTTCTTAAAGTGGCCAAATGTA	1680
Qy	1681	AATCTTAGAAACTATAAGATGAACAGTTTACACAGTCTTGTACGAGACTACTTTATTTT	1740
Db	1681	AATCTTAGAAACTATAAGATGAACAGTTTACACAGTCTTGTACGAGACTACTTTATTTT	1740
Qy	1741	TACAAGCCGACGAGTAAATTTATATGCCAAGCTGGATCTGGATTTTCCAGGCCAAACAG	1800
Db	1741	TACAAGCCGACGAGTAAATTTATATGCCAAGCTGGATCTGGATTTTCCAGGCCAAACAG	1800
Qy	1801	CTCAGGTTGTAGGTCGCCAGTTTACAGAAATTTCTTTGAATCTGAGAGGATGGGCAA	1860
Db	1801	CTCAGGTTGTAGGTCGCCAGTTTACAGAAATTTCTTTGAATCTGAGAGGATGGGCAA	1860
Qy	1861	GGCTACTTTAGAAAGTCTAGTAAAGGATATTTGTTCAAGTGCCTCAATGCTTCAATG	1920
Db	1861	GGCTACTTTAGAAAGTCTAGTAAAGGATATTTGTTCAAGTGCCTCAATGCTTCAATG	1920
Qy	1921	AAACCCGAAAGAGCTTTCAAAATATGTTTATTCACACCCCTTACTAGTCAACACTACTTT	1980
Db	1921	AAACCCGAAAGAGCTTTCAAAATATGTTTATTCACACCCCTTACTAGTCAACACTACTTT	1980
Qy	1981	TTATTTATTTGGAACAACCTTCTGCCACCTCATGGAGTTAAATATGCTGGAAAATGCACT	2040
Db	1981	TTATTTATTTGGAACAACCTTCTGCCACCTCATGGAGTTAAATATGCTGGAAAATGCACT	2040
Qy	2041	GTAATTTAGTCTCTCTTAATCTTTGCTCTGAAAACCAACATCACTTGTCTAAACCTT	2100
Db	2041	GTAATTTAGTCTCTCTTAATCTTTGCTCTGAAAACCAACATCACTTGTCTAAACCTT	2100
Qy	2101	ACTGTTTCTAGCTTGGACTATAGCAGAGATGGATTTGGCTAGAGTCACTCTTCCAAAAAT	2160
Db	2101	ACTGTTTCTAGCTTGGACTATAGCAGAGATGGATTTGGCTAGAGTCACTCTTCCAAAAAT	2160
Qy	2161	TTAACTGACGCTACTGATGCTTCAATTCAGATGAAAACGAGCTTATCCCACTTGGAGACA	

Db	3481	ACTTCATTATTATGGGAAATAAGCACTTTGAAGACACTGGTAGTACACCAAGCATTTGGAGAA	3541
Qy	3541	AATGACTTAAATTTACCAAGAAATTTTGGTACAGAGAAATCACAGAGAAATATACAGCCGA	3600
Db	3541	AATGACTTAAATTTACCAAGAAATTTTGGTACAGAGAAATCACAGAGAAATATACAGCCGA	3600
Qy	3601	GAGAGGTTAGTAGAGAAAGTTTCAACGAGCTTCATATGAAGATACGTAGCCAAAGTTTC	3660
Db	3601	GAGAGGTTAGTAGAGAAAGTTTCAACGAGCTTCATATGAAGATACGTAGCCAAAGTTTC	3660
Qy	3661	AATACAGACACTACAAAGTGGCATAAGTTCTAATGAGCTCAAGTCTTTCACGAGAGACA	3720
Db	3661	AATACAGACACTACAAAGTGGCATAAGTTCTAATGAGCTCAAGTCTTTCACGAGAGACA	3720
Qy	3721	GTAGGTGTAGATGCTACAACTATGAGACACAGACTGTGGAAGCATGAGTACTGTGGTAAAGT	3780
Db	3721	GTAGGTGTAGATGCTACAACTATGAGACACAGACTGTGGAAGCATGAGTACTGTGGTAAAGT	3780
Qy	3781	ACTAAAACTATTAAAGACAAGCCACTATTTGACGCCACAGCTCTAACCATCTGTCTCTCTCC	3840
Db	3781	ACTAAAACTATTAAAGACAAGCCACTATTTGACGCCACAGCTCTAACCATCTGTCTCTCTCC	3840
Qy	3841	AAATCAAAATTCGGTGCCCTGAGTTCAGCAAAAGTCTCTCTATAGGCTTCTTAGAAGACA	3900
Db	3841	AAATCAAAATTCGGTGCCCTGAGTTCAGCAAAAGTCTCTCTATAGGCTTCTTAGAAGACA	3900
Qy	3901	CAGTCCCTTAAAGCAGCACCTCTATTGTGTACAAATTAAGTCTAGCAGATTCGTAACTTTAGT	3960
Db	3901	CAGTCCCTTAAAGCAGCACCTCTATTGTGTACAAATTAAGTCTAGCAGATTCGTAACTTTAGT	3960
Qy	3961	TACACAAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAAGCTACAGCAACAAAGA	4020
Db	3961	TACACAAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAAGCTACAGCAACAAAGA	4020
Qy	4021	ATGCATCCATCTTATCTCACTCTGAAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTT	4080
Db	4021	ATGCATCCATCTTATCTCACTCTGAAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTT	4080
Qy	4081	ACTGATACCATCACCATGAAGGCCACAGTTTTTGGTCCAGATTTAACACCAAGCAGGTTC	4140
Db	4081	ACTGATACCATCACCATGAAGGCCACAGTTTTTGGTCCAGATTTAACACCAAGCAGGTTC	4140
Qy	4141	ATGAAGCCCTTAAGTTATGCATCATTAGATAAAGAAAGATTTATTGAGTCCCTATTAATCAA	4200
Db	4141	ATGAAGCCCTTAAGTTATGCATCATTAGATAAAGAAAGATTTATTGAGTCCCTATTAATCAA	4200
Qy	4201	AATACCCGTCAACGATCTTCTCAGTCCGCTCAGTGTCCAGTGCCACATATGGGGT	4260
Db	4201	AATACCCGTCAACGATCTTCTCAGTCCGCTCAGTGTCCAGTGCCACATATGGGGT	4260
Qy	4261	TCAGATGATTAATTTGTTCTCTCCGGTGGATATAAATCATATATTTCAGGTAAAG	4320
Db	4261	TCAGATGATTAATTTGTTCTCTCCGGTGGATATAAATCATATATTTCAGGTAAAG	4320
Qy	4321	GATATTTCCCTATTTCAGACAAAAACATACCAACCATGATGATCGAGTGCAAGAGCA	4380
Db	4321	GATATTTCCCTATTTCAGACAAAAACATACCAACCATGATGATCGAGTGCAAGAGCA	4380
Qy	4381	TTTGCCCATGATCGAGAGGTCTTCCATCTGGAACCTGGAGGTCTGTGTAAAAATTTCTTTT	4440
Db	4381	TTTGCCCATGATCGAGAGGTCTTCCATCTGGAACCTGGAGGTCTGTGTAAAAATTTCTTTT	4440
Qy	4441	CACCTGTCACGACAGAGATGAGTCTTACGGAAATATGAATTTCAATCCATTCAGATGCC	4500
Db	4441	CACCTGTCACGACAGAGATGAGTCTTACGGAAATATGAATTTCAATCCATTCAGATGCC	4500
Qy	4501	TCTCTGTTTTTGAAGAGTACAGAAAGTCTGGACTACAGGAAATACAGATGATACTGC	4560
Db	4501	TCTCTGTTTTTGAAGAGTACAGAAAGTCTGGACTACAGGAAATACAGATGATACTGC	4560
Qy	4561	CTTTATGTGTCGTATGTAAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCAATA	4620
Db	4561	CTTTATGTGTCGTATGTAAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCAATA	4620

QY 4621 TGTAGTCATTGAGACTTTCAAGATATTCATATTTCTGATTGGTGTGAGCAGACTATCCAT 4680
DB |||||
QY 4621 TGTAGTCATTGAGACTTTCAAGATATTCATATTTCTGATTGGTGTGAGCAGACTATCCAT 4680
DB |||||
QY 4681 AATCCTTTAGAGTGGTTCCTCTAAGATTTTCGGGATTTCTGGATGCACTGATGGGTG 4740
DB |||||
QY 4681 AATCCTTTAGAGTGGTTCCTCTAAGATTTTCGGGATTTCTGGATGCACTGATGGGTG 4740
DB |||||
QY 4741 TCTCAAGAAGGCTCAGCTAGCAGCACCACCAAGACAGAAATTTGTTACTAGGTGTTAAACA 4800
DB |||||
QY 4741 TCTCAAGAAGGCTCAGCTAGCAGCACCACCAAGACAGAAATTTGTTACTAGGTGTTAAACA 4800
DB |||||
QY 4801 ATTCCAGATCATACCAATGTCGGTATATCTCTCGCAAGAGTTCCTAAGATTAGTC 4860
DB |||||
QY 4801 ATTCCAGATCATACCAATGTCGGTATATCTCTCGCAAGAGTTCCTAAGATTAGTC 4860
DB |||||
QY 4861 ATTAATTTGAGTGTTCAGTTTCAACTAAATGTCTGAGACTGGGCTTTTAAACAATTAAG 4920
DB |||||
QY 4861 ATTAATTTGAGTGTTCAGTTTCAACTAAATGTCTGAGACTGGGCTTTTAAACAATTAAG 4920
DB |||||
QY 4921 GAGAAGTATCTCAACATTTGATGACATATGCTTTACTCTGAGGTTTCCCAATTTGCTG 4980
DB |||||
QY 4921 GAGAAGTATCTCAACATTTGATGACATATGCTTTACTCTGAGGTTTCCCAATTTGCTG 4980
DB |||||
QY 4981 TCACACTGCACATTCAGACTTCGGTGTGGAGTTCATACAGAAATTTATTTCAAGATGTA 5040
DB |||||
QY 4981 TCACACTGCACATTCAGACTTCGGTGTGGAGTTCATACAGAAATTTATTTCAAGATGTA 5040
DB |||||
QY 5041 CAGTTTCTACAAATGTCATGAAGAGCAGAGGCTGTGTGGCAACACCAACCAAGCAACCT 5100
DB |||||
QY 5041 CAGTTTCTACAAATGTCATGAAGAGCAGAGGCTGTGTGGCAACACCAACCAAGCAACCT 5100
DB |||||
QY 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127
DB |||||
QY 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127
DB |||||

RESULT 2

Q0869395
LOCUS C0869395 5662 bp DNA linear PAT 13-SBP-2004
DEFINITION Sequence 1 from Patent WO2004074448.
ACCESSION C0869395
VERSION C0869395.1 GI:51999314
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sabatini D.M., Kim D.H. and Sarbasov D.D. ✓
TITLE MTOR KINASE-ASSOCIATED PROTEINS
JOURNAL Patent: WO 2004074448-A 1 02-SEP-2004;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES
source
1. .5662
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5127; DB 6; Length 5662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCGCGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC 60
DB |||||
QY 28 ATGCGCGCGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC 87
DB |||||
QY 61 AGCGGCGAGGAGAACGTCCTCGCTGGATCTGACCCGAGAACCTTCTGATTAACCTTAAGAGAG 120
DB |||||
QY 88 AGCGGCGAGGAGAACGTCCTCGCTGGATCTGACCCGAGAACCTTCTGATTAACCTTAAGAGAG 147
DB |||||

QY 121 ATTCTCCAAAATGTGGCCAGATTGCGAGGATCAAAATATGAGAAAGCTAGGCCATCTG 180
DB |||||
QY 148 ATTCTCCAAAATGTGGCCAGATTGCGAGGATCAAAATATGAGAAAGCTAGGCCATCTG 207
DB |||||
QY 181 AATAAATTCTAAGCTTCTTTGTGATATGCGCACAGTGAGAAAGAACTGGGCTTTTAC 240
DB |||||
QY 208 AATAAATTCTAAGCTTCTTTGTGATATGCGCACAGTGAGAAAGAACTGGGCTTTTAC 267
DB |||||
QY 241 TATGAGATATCATTAATTTTGGGTTAGCTTTTAAATGAAAGAAAGAGTGGCA 300
DB |||||
QY 268 TATGAGATATCATTAATTTTGGGTTAGCTTTTAAATGAAAGAAAGAGTGGCA 327
DB |||||
QY 301 GCAGCAGGCTCAGAGGCTTCGATATCTCATCCAGACTCCAGTATTTCTCCAGAGGTTG 360
DB |||||
QY 328 GCAGCAGGCTCAGAGGCTTCGATATCTCATCCAGACTCCAGTATTTCTCCAGAGGTTG 387
DB |||||
QY 361 CTAAATTTGAAAGTGGACTTATTAATAGTGGTGCATTGACATACAAAGAGCAACGAG 420
DB |||||
QY 388 CTAAATTTGAAAGTGGACTTATTAATAGTGGTGCATTGACATACAAAGAGCAACGAG 447
DB |||||
QY 421 GTAGAGGACACACAGCCTTCGATTTAGTTCAGAAAGATGATTACTGTGAATCTTCTTCTG 480
DB |||||
QY 448 GTAGAGGACACACAGCCTTCGATTTAGTTCAGAAAGATGATTACTGTGAATCTTCTTCTG 507
DB |||||
QY 481 TTTCTTAGTTCTGTCACCAACTCATTAATGTCAGTTGGAAATGATGGAATTTCAAGAAAGA 540
DB |||||
QY 508 TTTCTTAGTTCTGTCACCAACTCATTAATGTCAGTTGGAAATGATGGAATTTCAAGAAAGA 567
DB |||||
QY 541 GACAGAAATGTCGAGCATGTCATTCTGTGAACTAGCACTTCAGAAATCCAGAG 600
DB |||||
QY 568 GACAGAAATGTCGAGCATGTCATTCTGTGAACTAGCACTTCAGAAATCCAGAG 627
DB |||||
QY 601 GTGGTGGCCCTTCGAGGAGCTTAAACACCATTTGAAAATGTAATGTCATTTGCCAATTA 660
DB |||||
QY 628 GTGGTGGCCCTTCGAGGAGCTTAAACACCATTTGAAAATGTAATGTCATTTGCCAATTA 687
DB |||||
QY 661 AGTCGAATAAATGAGGCCCTTAATTAACAATTTTGCACCTTCTTAATCATCCAAAGACT 720
DB |||||
QY 688 AGTCGAATAAATGAGGCCCTTAATTAACAATTTTGCACCTTCTTAATCATCCAAAGACT 747
DB |||||
QY 721 CGGCAGTATGTGCGAGCTGATGTAGAAATTTAGAGAAATTTTAGCACCTTACTGATTTT 780
DB |||||
QY 748 CGGCAGTATGTGCGAGCTGATGTAGAAATTTAGAGAAATTTTAGCACCTTACTGATTTT 807
DB |||||
QY 781 CACTACAGCATATGTCAGATACAGCTGAGCAGCAGCTCAAGAGACAGAGAGACCA 840
DB |||||
QY 808 CACTACAGCATATGTCAGATACAGCTGAGCAGCAGCTCAAGAGACAGAGAGACCA 867
DB |||||
QY 841 TTTCTAGCCAGTAAATGGAATTCATAGCAATTCAGATCCGATCATGGGAGGTATTATTAAT 900
DB |||||
QY 868 TTTCTAGCCAGTAAATGGAATTCATAGCAATTCAGATCCGATCATGGGAGGTATTATTAAT 927
DB |||||
QY 901 TTATGTAAACCTGGAAATTTCTGGATCCAGTCTCTTAATAGGAGTACTTTGCAATACCAAT 960
DB |||||
QY 928 TTATGTAAACCTGGAAATTTCTGGATCCAGTCTCTTAATAGGAGTACTTTGCAATACCAAT 987
DB |||||
QY 961 ATGGAATAAGCGGAGGTCTCTGAGGCTTTTATGATATTTTCTCTCTTCTTCTTACCT 1020
DB |||||
QY 988 ATGGAATAAGCGGAGGTCTCTGAGGCTTTTATGATATTTTCTCTTCTTCTTACCT 1047
DB |||||
QY 1021 GTTGTGACTGAGGAGTTTCATAGAAGCACTACTCAGTGTAGATCCAGGGAGGTTCGAAGAC 1080
DB |||||
QY 1048 GTTGTGACTGAGGAGTTTCATAGAAGCACTACTCAGTGTAGATCCAGGGAGGTTCGAAGAC 1107
DB |||||
QY 1081 AGTTGAGGCTTTGAGATGGCTTTTGGCAGCTGAGGCAAAACCTATTTCTTCTCATCGT 1140
DB |||||
QY 1108 AGTTGAGGCTTTGAGATGGCTTTTGGCAGCTGAGGCAAAACCTATTTCTTCTCATCGT 1167
DB |||||
QY 1141 GCCAGATCCAGGCCAGCCTCATGTAATTTATTTGGCACTGATCTCTCTGATTTAT 1200
DB |||||
QY 1168 GCCAGATCCAGGCCAGCCTCATGTAATTTATTTGGCACTGATCTCTCTGATTTAT 1227
DB |||||
QY 1201 CGTAATGGACTTTTAGAGGCTCTAGTTGAAGTGATGATAACAAACAGTGATGATCATATCTCA 1260
DB |||||

Db 1228 GCTAATGACATTTTAGAGGCTCTAGTTGAAGGATGATAACAAACAGTGATGATCATCTCA 1287
Qy 1261 GTTAGAGCTACCATCTCTTTAGGAGAGCTTTTACATATGGCAAAACACAATTTCTTCTCAT 1320
Db 1288 GTTAGAGCTACCATCTCTTTAGGAGAGCTTTTACATATGGCAAAACACAATTTCTTCTCAT 1347
Qy 1321 TCACATAGCCATCATTTACATCTGCTGTCGCAACCCCTTAATGAATATGGCTGCAATCTTTGAT 1380
Db 1348 TCACATAGCCATCATTTACATCTGCTGTCGCAACCCCTTAATGAATATGGCTGCAATCTTTGAT 1407
Qy 1381 ATCCCCAGGAAAGAGAGCTCGAGCCAGTCGAGCTTGAAGCTGTTTAAAGCTTCCAT 1440
Db 1408 ATCCCCAGGAAAGAGAGCTCGAGCCAGTCGAGCTTGAAGCTGTTTAAAGCTTCCAT 1467
Qy 1441 GAAATGAAGAAACGAGGACCTTAAGCCTTATAGTCTTCAATTTAGACCAATTAATTCAGAAA 1500
Db 1468 GAAATGAAGAAACGAGGACCTTAAGCCTTATAGTCTTCAATTTAGACCAATTAATTCAGAAA 1527
Qy 1501 GCAATGCAACACACCCAGAAACGGGATCAGTATCTCGAGTTCAGAAAGATATATTTATC 1560
Db 1528 GCAATGCAACACACCCAGAAACGGGATCAGTATCTCGAGTTCAGAAAGATATATTTATC 1587
Qy 1561 CTTAAGGATACAGAGGAGCTCTTTAATTAACCTTAGAGATAGCCAGTCTTCAACAT 1620
Db 1588 CTTAAGGATACAGAGGAGCTCTTTAATTAACCTTAGAGATAGCCAGTCTTCAACAT 1647
Qy 1621 AAAGAGATCTTGAATGGAATCTTATAGGACCAATCTTAAGTGCCCAAAATGTA 1680
Db 1648 AAAGAGATCTTGAATGGAATCTTATAGGACCAATCTTAAGTGCCCAAAATGTA 1707
Qy 1681 AATCTAGAAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTATTTATTTT 1740
Db 1708 AATCTAGAAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTATTTATTTT 1767
Qy 1741 TACAGCCAGCAGTAAATATATATGCCAACCTGATCTGGATTTTGGCAAGGCCAAACAG 1800
Db 1768 TACAGCCAGCAGTAAATATATATGCCAACCTGATCTGGATTTTGGCAAGGCCAAACAG 1827
Qy 1801 CTACAGGTTGAGGTTGCGAGTTTACAGAAATTTCTTTGAAATCTGAAGAGATGGGCAA 1860
Db 1828 CTACAGGTTGAGGTTGCGAGTTTACAGAAATTTCTTTGAAATCTGAAGAGATGGGCAA 1887
Qy 1861 GGCTACTTAGAAGATCTAGTAAAGATATTTGTCAGTGGCTCAATGCTTCACTCGGAATG 1920
Db 1888 GGCTACTTAGAAGATCTAGTAAAGATATTTGTCAGTGGCTCAATGCTTCACTCGGAATG 1947
Qy 1921 AAACCCGAAAGAGCTCTTCAAAATTAATGGTTTATTTGACCAACCTTTAGTCAACACTATTT 1980
Db 1948 AAACCCGAAAGAGCTCTTCAAAATTAATGGTTTATTTGACCAACCTTTAGTCAACACTATTT 2007
Qy 1981 TTATTTATTTGAACACATTTCTTTCACCCCTCATCGAGTTTAAATGCTGGAAAATGCAAT 2040
Db 2008 TTATTTATTTGAACACATTTCTTTCACCCCTCATCGAGTTTAAATGCTGGAAAATGCAAT 2067
Qy 2041 GTATTTCAAGTCTCTCTTAATCTTTTGTCTCTTGAAGAACCAAGATCACTTGTCTAAACTT 2100
Db 2068 GTATTTCAAGTCTCTCTTAATCTTTTGTCTCTTGAAGAACCAAGATCACTTGTCTAAACTT 2127
Qy 2101 ACTGTTTCTAGCTTGGACTATAGCAGATGGATGGCTAGAGTCACTCTTCCAAATTT 2160
Db 2128 ACTGTTTCTAGCTTGGACTATAGCAGATGGATGGCTAGAGTCACTCTTCCAAATTT 2187
Qy 2161 TTAATCTCAGTCTAGTCTGAGACTCTATGCAACAAACATTTAAGGGTATTTATG 2220
Db 2188 TTAATCTCAGTCTAGTCTGAGACTCTATGCAACAAACATTTAAGGGTATTTATG 2247
Qy 2221 AGAGCTAATGTTGAATCTTTTAAATTTGGGAAATTCAGTGTGTAGTGACCCAGCTACAT 2280
Db 2248 AGAGCTAATGTTGAATCTTTTAAATTTGGGAAATTCAGTGTGTAGTGACCCAGCTACAT 2307
Qy 2281 GATAAAACAAACGATTTCTCTGAAGCTTTGATATCTCGATGAAGCATGTGAAGAC 2340

Db 2308 GATAAAACAAACGATTTCTCTGAAGCTCTTGATATCTCGATGAAGCATGTGAAGAC 2367
Qy 2341 AAGGCCAATCTTCAATGCTCTCATTTAGATGAACCCAGGTTTATCCACCTTTGGAGACAAG 2400
Db 2368 AAGGCCAATCTTCAATGCTCTCATTTAGATGAACCCAGGTTTATCCACCTTTGGAGACAAG 2427
Qy 2401 GGTTTGTCTCTCTGCTGAGATTTCTCTCAATCCAAAAGGATTTTCTTCTCTGAATGAA 2460
Db 2428 GGTTTGTCTCTCTGCTGAGATTTCTCTCAATCCAAAAGGATTTTCTTCTCTGAATGAA 2487
Qy 2461 AGAGGTTATGAGCAAAACAAATTCGAAAAGTGGCACAGGGAATACAACTCCAAATATGTT 2520
Db 2488 AGAGGTTATGAGCAAAACAAATTCGAAAAGTGGCACAGGGAATACAACTCCAAATATGTT 2547
Qy 2521 GACTTTGATGAGGAAACAACTCAATGAAGCACTTACTACTTACCGGAAGCCCTGTTGATGTT 2580
Db 2548 GACTTTGATGAGGAAACAACTCAATGAAGCACTTACTACTTACCGGAAGCCCTGTTGATGTT 2607
Qy 2581 GATAAATATGTTCTGCGAGTAAACCAAGATTTACAGCGTCTCTCACTCTACCTGCCTATA 2640
Db 2608 GATAAATATGTTCTGCGAGTAAACCAAGATTTACAGCGTCTCTCACTCTACCTGCCTATA 2667
Qy 2641 CACCTTTATGGAACAACTAGTACACCAATAAAACAGGCTGCCATTTTGTGGAAAGTACAGAT 2700
Db 2668 CACCTTTATGGAACAACTAGTACACCAATAAAACAGGCTGCCATTTTGTGGAAAGTACAGAT 2727
Qy 2701 ATTATACAGAACTCTCTGTAATGTTTCTGACACAGATTTGGATTAAGTGGGAAGAAATTT 2760
Db 2728 ATTATACAGAACTCTCTGTAATGTTTCTGACACAGATTTGGATTAAGTGGGAAGAAATTT 2787
Qy 2761 AAAAACTGAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAAAATTTGGGCTCTC 2820
Db 2788 AAAAACTGAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAAAATTTGGGCTCTC 2847
Qy 2821 AATTGCTACAGGAAGAAACGGTATTCAGATATATACTAAAACTTGCAGAAACAGTGTGAA 2880
Db 2848 AATTGCTACAGGAAGAAACGGTATTCAGATATATACTAAAACTTGCAGAAACAGTGTGAA 2907
Qy 2881 GTTCTTTCCATCAGAGGACCTGTGTATATGTTCTGCGGCTCATAGCTTAAACCAACAA 2940
Db 2908 GTTCTTTCCATCAGAGGACCTGTGTATATGTTCTGCGGCTCATAGCTTAAACCAACAA 2967
Qy 2941 GGCTGTGATATTTCAAAATGTCACAACTGGGATGCTGTGAGGCAATAGTCGCAAAACATCTG 3000
Db 2968 GGCTGTGATATTTCAAAATGTCACAACTGGGATGCTGTGAGGCAATAGTCGCAAAACATCTG 3027
Qy 3001 TGGCCAGTGTTCAGATGATGTGGAAACAACTCTGTAATGAACTTTTCACTATCCCAAGC 3060
Db 3028 TGGCCAGTGTTCAGATGATGTGGAAACAACTCTGTAATGAACTTTTCACTATCCCAAGC 3087
Qy 3061 ACTCTAAGTTTGAACCTCGGAGTCAACCCAGCTCTAGACATATAGTGAAGTGAATCTGTG 3120
Db 3088 ACTCTAAGTTTGAACCTCGGAGTCAACCCAGCTCTAGACATATAGTGAAGTGAATCTGTG 3147
Qy 3121 CCATCCAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
Db 3148 CCATCCAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3207
Qy 3181 TTCTTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCATAAAG 3240
Db 3208 TTCTTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCATAAAG 3267
Qy 3241 GATAAAATTTCAATTTCTTTCTTTCTTTAGTAACTTTGAGAGATCTGATCTTCTTAAT 3300
Db 3268 GATAAAATTTCAATTTCTTTCTTTCTTTAGTAACTTTGAGAGATCTGATCTTCTTAAT 3327
Qy 3301 TCGCTTACTTTGCTTAAACAAAAACATCTGATAGCAGTATCCAAAAGGAGGAAATTA 3360
Db 3328 TCGCTTACTTTGCTTAAACAAAAACATCTGATAGCAGTATCCAAAAGGAGGAAATTA 3387
Qy 3361 TCATCTGAAGATGAAGCAACAGCCGAATCAGAACACTTACGGAGCCAGTGTGAT 3420
Db 3388 TCATCTGAAGATGAAGCAACAGCCGAATCAGAACACTTACGGAGCCAGTGTGAT 3447

3421 TTTAATCATAGTATGATTTTACACCATATCCATCTGTACAGAAACATTAACAATTAGAG 3480
3448 TTTAATCATAGTATGATTTTACACCATATCCATCTGTACAGAAACATTAACAATTAGAG 3507
3481 ACTTCATTTATGGGGAATTAAGCACATCTGAAGACACTGTGTAGTACACCAAGCAATTTGGAGAA 3540
3508 ACTTCATTTATGGGGAATTAAGCACATCTGAAGACACTGTGTAGTACACCAAGCAATTTGGAGAA 3567
3541 AATGACTTAAATTTACCAAGAAATTTTGGTACAGAGAAATCACAGAGAAATTAACAAGCCGA 3600
3568 AATGACTTAAATTTACCAAGAAATTTTGGTACAGAGAAATCACAGAGAAATTAACAAGCCGA 3627
3601 GAGAGGTATAGTATGAGAAATTTCAACGAGCTCACATATGAAGATAGTACGAAAGTTTC 3660
3628 GAGAGGTATAGTATGAGAAATTTCAACGAGCTCACATATGAAGATAGTACGAAAGTTTC 3687
3661 AATACAGACACTACCAAGTGGCATAAGTTCAATGAGCTCAAGTCTTTCACAGAGAGACA 3720
3688 AATACAGACACTACCAAGTGGCATAAGTTCAATGAGCTCAAGTCTTTCACAGAGAGACA 3747
3721 GTAGGTATAGTATGAGAAATTTTGGTACAGAGAAATTAACAAGCCGA 3780
3748 GTAGGTATAGTATGAGAAATTTTGGTACAGAGAAATTAACAAGCCGA 3807
3781 ACTTAAACTATTAAGACAGCACTATTTGAGCGCCACAGCTCAACCATCTCTCTCTCC 3840
3808 ACTTAAACTATTAAGACAGCACTATTTGAGCGCCACAGCTCAACCATCTCTCTCTCC 3867
3841 AATCAAAATTCGGTGTCCCTGGTGTCCCTCCAGGTTCTTCTCATACGCTTCTTCTAGAGAGACA 3900
3868 AATCAAAATTCGGTGTCCCTGGTGTCCCTCCAGGTTCTTCTCATACGCTTCTTCTAGAGAGACA 3927
3901 CAGTCCCTTAAAGCACTCTTATGCTACATTAATAAGTCTAGCAGATTTGTAATTTAGT 3960
3928 CAGTCCCTTAAAGCACTCTTATGCTACATTAATAAGTCTAGCAGATTTGTAATTTAGT 3987
3961 TACACAAGTCTAGAGATGCTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA 4020
3988 TACACAAGTCTAGAGATGCTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA 4047
4021 ATGCATCCATCTTATCTCACTGAGGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4080
4048 ATGCATCCATCTTATCTCACTGAGGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4107
4081 ACTGATACCATCACCATGAGGCACTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA 4140
4108 ACTGATACCATCACCATGAGGCACTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA 4167
4141 ATGAAAGCCCTTAAGTTATGATCATCATTTAGATAAAGAGATTTTATGAGTCTTATTAATCAA 4200
4168 ATGAAAGCCCTTAAGTTATGATCATCATTTAGATAAAGAGATTTTATGAGTCTTATTAATCAA 4227
4201 AATACCCCTGCAACGATCTTCTCAGTGGCGTCCATGGTTCAGTGCCCAATATGGGGT 4260
4228 AATACCCCTGCAACGATCTTCTCAGTGGCGTCCATGGTTCAGTGCCCAATATGGGGT 4287
4261 TCAGATGATTAACATTTGCTTCTTCTCCCGGTGGATATAAATGATATATTTCCAGGTAAAG 4320
4288 TCAGATGATTAACATTTGCTTCTTCTCCCGGTGGATATAAATGATATATTTCCAGGTAAAG 4347
4321 GATATTCCTTATTTTACAGCAAAACATACACCAACCATGATGATGATGATGATGATGATGATGAT 4380
4348 GATATTCCTTATTTTACAGCAAAACATACACCAACCATGATGATGATGATGATGATGATGATGAT 4407
4381 TTTGCCCATGATGAGGAGTCTTCCATCTGGAACCTGGAGGCTCTGTAAATAATCTTTT 4440
4408 TTTGCCCATGATGAGGAGTCTTCCATCTGGAACCTGGAGGCTCTGTAAATAATCTTTT 4467
4441 CACTTGTACAGAGAGATGAGTCTTACCGAAATTAATGAATTCATCCATTCAGATGCC 4500
4468 CACTTGTACAGAGAGATGAGTCTTACCGAAATTAATGAATTCATCCATTCAGATGCC 4527

4501 TCTCTGTTTTAGAAAGTACAGAAAGACACTGAGCTACAGGAAACATACAGATGATAACTGC 4560
4528 TCTCTGTTTTAGAAAGTACAGAAAGACACTGAGCTACAGGAAACATACAGATGATAACTGC 4587
4561 CTTTATTTGCTGCTGATTTGAAATTTCTGGGTTTTCAGCCCAACCACTGAGTGCAATA 4620
4588 CTTTATTTGCTGCTGATTTGAAATTTCTGGGTTTTCAGCCCAACCACTGAGTGCAATA 4647
4621 TGTAGTCAATTCAGACCTTTCAAGATATTCATATTTCTGATTTGTTGAGCAGACTATCCAT 4680
4648 TGTAGTCAATTCAGACCTTTCAAGATATTCATATTTCTGATTTGTTGAGCAGACTATCCAT 4707
4681 AATCCTTTTGAAGTGGTTCCTCTAAAGTTTTCGGGATTTCTGGATGAGTGAGGGTG 4740
4708 AATCCTTTTGAAGTGGTTCCTCTAAAGTTTTCGGGATTTCTGGATGAGTGAGGGTG 4767
4741 TCTCAGAGAGCTCAGCTAGCAGCACAAGACCAAGACCAAGAAATTTGTTACTAGGTGTTAAACA 4800
4768 TCTCAGAGAGCTCAGCTAGCAGCACAAGACCAAGACCAAGAAATTTGTTACTAGGTGTTAAACA 4827
4801 ATTCCAGATGATACACCAATGTGCGGTATATCTCTTCGCAAGAGATTTCTAAGATTAGTC 4860
4828 ATTCCAGATGATACACCAATGTGCGGTATATCTCTTCGCAAGAGATTTCTAAGATTAGTC 4887
4861 ATTAATTTGAGTGTAGTTTCAACTAAATGTCTAAGATGCTGGGCTTTTAAACAATTAAAG 4920
4888 ATTAATTTGAGTGTAGTTTCAACTAAATGTCTAAGATGCTGGGCTTTTAAACAATTAAAG 4947
4921 GAGAGATATCTCAAAACATTTGATGACATATGCTTTACTCTGAGGTTTCCCATTTGCTG 4980
4948 GAGAGATATCTCAAAACATTTGATGACATATGCTTTACTCTGAGGTTTCCCATTTGCTG 5007
4981 TCACACTGCATTCAGACTTCCCGTGTGCGAGGTTTCATACAGAAATTTTCAAGATGTA 5040
5008 TCACACTGCATTCAGACTTCCCGTGTGCGAGGTTTCATACAGAAATTTTCAAGATGTA 5067
5041 CAGTTTCTCAAAATGCATGAAAGACAGAGCTGTGTGGCAACACCAACCAAGCAACT 5100
5068 CAGTTTCTCAAAATGCATGAAAGACAGAGCTGTGTGGCAACACCAACCAAGCAACT 5127
5101 ATAGTTGATACATCTGCTGAATCCCTGA 5127
5128 ATAGTTGATACATCTGCTGAATCCCTGA 5154

RESULT 3
AY515854
LOCUS
DEFINITION Homo sapiens rapamycin insensitive companion of mTOR mRNA, complete cds.
ACCESSION AY515854
VERSION AY515854.1 GI:46093885
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5662)
AUTHORS Dos D., Ali, S.M., Kim, D.H., Guertin, D.A., Latek, R.R., Erdjument-Bronage, H., Tempst, P., and Sabatini, D.M.
TITLE Rictor, a Novel Binding Partner of mTOR, Defines a Rapamycin-Insensitive and Raptor-Independent Pathway that Regulates the Cytoskeleton
JOURNAL Curre. Biol. 14 (14), 1296-1302 (2004)
PUBMED 15268862

REFERENCE 2 (bases 1 to 5662)
AUTHORS Sabatini, D.M., Sarbassov, D.D., Ali, S.M., Kim, D.-H. and Guertin, D.A.
TITLE Direct Submmission
JOURNAL Submitted (02-JAN-2004) Sabatini Laboratory, Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

FEATURES
source Location/Qualifiers
1. .5662

CDS	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	28. .5154	
	/note="mTOR binding protein; rictor"	
	/codon_start=1	
	/product="rapamycin insensitive companion of mTOR"	
	/protein_id="AAS79796.1"	
	/db_xref="GI:46093886"	
	/translation="MAAIGRGRSLKRLNRVRGRNDGSENVPLDLITREPSDNLREILQN VARLQGVNMRKGLHNLNFKLLCDIGHSBEKGFHVEDIIICLRALLNEAEVRAA GLRALVLIQDSSTLOKVLKLVYDLIARCIDIQOSNEVETQALRLVRKMTVNASL FPSSVTSLIANGNDLQEDRMVRACIIICELAQNPVVALRGRLITLKNVIDC QLSRINBALITLHLNLHPKTYQVRADYVELERILAPYDFHYRHSPTDAEQLKED REARFLASKMGIITAFRSWAGIINLCPGNSGQSILGVLCIPNMEIRGLLEVLDI FRPLPVVTEFTEALSDVPGRFQDSWRSLSDGFVAEAKTILPHRARSPRLMDNVL ALILSAFIRNGLEGLVEVITNSDDHISVRATILIGELHWMANTILPHSHSHLCLP TLNMAASFDIPEKRLASALNCLRFHEMKRGPYPYSLHDHIIQKAIATHQKR DEYLVRQDILFKDTEBALINLRDSQVLQHNLEWNNLIGITLKNPNNLRNYK QDLHRVFLNLYFPKSLYANLDFPAKAKQLTVVGCQFTFELLESDEEQGYLE DLVKDIQWLNASRKMPERSLQNNGLLITLSQHYFLFIGTSLCHPHGVKMLKCSVF QCLNLCSLKNQDHLKLVSSLDYSRDGLARVLISKLITAAQDCKRLYATKHLRVL RANVEFPNNGIELLTQLDKNKTISSEALDILDEACEDKANLHALIQMKPALSHLG DKGLLLLRPLSPKGSYLNERGYAKOLKXHEKREYNSKVDLIEBQLEALTYTRK PVGDNTRVRSNQRQLPHYLPILHYQLVHHTGCHLLEVNIIITELCNVRITPDL DWEEIKKASLWALNGISNWNGLNLQENVIPIDLKARQCEVLISIRGTCVTVL GLIAKTQGGDILKCHNWDVRSRKLWLPVDPDVEQLCNELSSIPSTLSLNSBST SRNSESPPSSMFILEDORFGSSSTSPFLDINEDTBTFYDRSGPIKDKNSPFPF ASSKLKNRILSLITLPNKKHRSDDPKGKLSSEKTSNRRITLTLEPSVDFNHSDD PTPISTVQKTLQLETSPGNKHIEDTGSTPSIGENDLKFTYNFETENHRENTSRRLV VESTSHMKIRSSQFNFTDTTIGSISSMSPSRETVGVDAITMDIDDCGSMSTVVSFK TIKTSHYLTQSNHLSLSKNSVSLVPPGSHTLPRRAQSLKAPSIATIKSLADCNFS YTSRDAFYATLKRLOQRMHPSLSSEALASPAKDVLPFTDITWMKANSFESRLTFS RFKMALSAYSLDKEDLILSPINQTLQRSSSVRSWVSATYGGSDDIIGLALPVDINDI FOVKDIPIFOTKNIPEHDDRGARAFADAGLPGSTGGLVKNFHLRQOMSLTEIMN SINSDASLFLESTEDQLEHTDDNCLYCVCIILGFQPSNQLSACSMSDFQDIPIYS DWCEQTHNLEVPVPSFGSCGSCVSGESASSTKELLILGVKTIIPDDTPMCR LRLREVLRLVINLSSVSTKCHETGLTIKERYPTQPDIDICLYSEVSHLLSHCTFLRP CRPFLOELFQDVQFLQHEBAEAVLATPPKQPIVDTSAS"	
277. .1062		
misc_feature	/note="Region: conserved region 1"	
misc_feature	1099. .1344	
misc_feature	/note="Region: conserved region 2"	
misc_feature	1432. .1521	
misc_feature	/note="Region: conserved region 3"	
misc_feature	1633. .1929	
misc_feature	/note="Region: conserved region 4"	
misc_feature	1996. .2082	
misc_feature	/note="Region: conserved region 5"	
misc_feature	2116. .2472	
misc_feature	/note="Region: conserved region 6"	
misc_feature	2497. .2607	
misc_feature	/note="Region: conserved region 7"	
misc_feature	2656. .2742	
misc_feature	/note="Region: conserved region 5"	
misc_feature	2794. .2874	
misc_feature	/note="Region: conserved region 5"	
misc_feature	2926. .3012	
misc_feature	/note="Region: conserved region 5"	
ORIGIN		
Query Match	100.0%; Score 5127; DB 9; Length 5662;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 5127; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGCGCGCGATCGCGCGCGCGCTCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC 60	
Db	28 ATGCGCGCGATCGCGCGCGCGCTCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC 87	
Qy	61 AGCGGCGAGGAGACGTCGCGCTGATCTCACCGGAGACCTTCGTGATACTTTAAGAGAG 120	
Db	88 AGCGGCGAGGAGAACGTCGCGCTGATCTCACCGGAGACCTTCGTGATACTTTAAGAGAG 147	

ORIGIN

Query Match	100.0%	Score	5127	DB	9	Length	5662
Best Local Similarity	100.0%	Pred. No.	0	Mismatches	0	Indels	0
Matches 5127	Conservative	0					
Qy	1	ATGCGCGGATCGCGCGCGCGCTCTCTGAGAACCTCCGAGTACGAGGCGGGAATGAC	60				
Db	28	ATGCGCGGATCGCGCGCGCGCTCTCTGAGAACCTCCGAGTACGAGGCGGGAATGAC	87				
Qy	61	AGCGCGGAGGAGACGTCCGCTGGATCTGACCCGAGAACCTTCTGATAACTTAAGAGAG	120				
Db	88	AGCGCGGAGGAGACGTCCGCTGGATCTGACCCGAGAACCTTCTGATAACTTAAGAGAG	147				

Db		1228	CGTAATGGAGCTTTTAGAGGGTCTAGTTGAAGTGATAACAAACAGTGGATGATCATATCTCA	1287
QY		1261	GTTAGAGCTACCAATCCTTTTAGAGAGCTTTTACATATGGGAAACAAATTTCTTCTCAT	1320
Db		1288	GTTAGAGCTACCAATCCTTTTAGAGAGCTTTTACATATGGGAAACAAATTTCTTCTCAT	1347
QY		1321	TCACATAGCCATCATTTACACTGCTTGGCCAAACCCCTAAATATGCTGCTCATCTTTTGAT	1380
Db		1348	TCACATAGCCATCATTTACACTGCTTGGCCAAACCCCTAAATATGCTGCTCATCTTTTGAT	1407
QY		1381	ATCCCCAAGGAAAGAGACTGGAGCCAGTGCAGCCCTTGAATGTTTAAAGCGCTTCCAT	1440
Db		1408	ATCCCCAAGGAAAGAGACTGGAGCCAGTGCAGCCCTTGAATGTTTAAAGCGCTTCCAT	1467
QY		1441	GAAATGAAGAACAGGAGACTTAAGCCCTTATAGTCTTTCAATTTAGACACATTTATTCAGAA	1500
Db		1468	GAAATGAAGAACAGGAGACTTAAGCCCTTATAGTCTTTCAATTTAGACACATTTATTCAGAA	1527
QY		1501	GCAATTTGCAACACACAGAAACGGGATCAGTATCTCCGAGTTTCAGAAAGATATATTATC	1560
Db		1528	GCAATTTGCAACACACAGAAACGGGATCAGTATCTCCGAGTTTCAGAAAGATATATTATC	1587
QY		1561	CTTAAAGGATACAGAGGAAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGTCTTCAACAT	1620
Db		1588	CTTAAAGGATACAGAGGAAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGTCTTCAACAT	1647
QY		1621	AAAGAAATCTTGAATGGAAATTTGGAATCTTATAGGAGCAATCTTAAAGTGGGCAAAATGTA	1680
Db		1648	AAAGAAATCTTGAATGGAAATTTGGAATCTTATAGGAGCAATCTTAAAGTGGGCAAAATGTA	1707
QY		1681	AATCTAAGAAACCTATAAAGATGAACAGTTTACACAGTTTGTACGAGTCTTCTTATTTT	1740
Db		1708	AATCTAAGAAACCTATAAAGATGAACAGTTTGTACGAGTCTTCTTATTTT	1767
QY		1741	TACAAGCCAGCAGTAATATATATGTCACCACTGGATCTGGATTTTGGCCAAAGGCCAAACAG	1800
Db		1768	TACAAGCCAGCAGTAATATATATGTCACCACTGGATCTGGATTTTGGCCAAAGGCCAAACAG	1827
QY		1801	CTCACGGTTGTAGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAAGAGATGGGCAA	1860
Db		1828	CTCACGGTTGTAGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAAGAGATGGGCAA	1887
QY		1861	GGCTACTTAGAAGATCTAGTAAAGGATATTTGTTCACTGAGTCTCAATCTTCAATCTGGAATG	1920
Db		1888	GGCTACTTAGAAGATCTAGTAAAGGATATTTGTTCACTGAGTCTCAATCTTCAATCTGGAATG	1947
QY		1921	AAACCCGAAAGAGTCTTCAAAATAATGGTTTATTCACCACTTGTAGTCAACACTACTTT	1980
Db		1948	AAACCCGAAAGAGTCTTCAAAATAATGGTTTATTCACCACTTGTAGTCAACACTACTTT	2007
QY		1981	TTATTTATTTGGAAACACTTTCTTGCCACCTCATGGAGTTTAAATGCTGGAATAATGTCAGT	2040
Db		2008	TTATTTATTTGGAAACACTTTCTTGCCACCTCATGGAGTTTAAATGCTGGAATAATGTCAGT	2067
QY		2041	GTAATTCAGTCTCCTTAATCTTCTGCTTCAAAACCAAGATCACATTTGCTTAAACCTT	2100
Db		2068	GTAATTCAGTCTCCTTAATCTTCTGCTTCAAAACCAAGATCACATTTGCTTAAACCTT	2127
QY		2101	ACTGTTTCTAGCTTGGACTATAGCAGATGGAATTTGGCTTAGAGTCACTCTTCCAAATTT	2160
Db		2128	ACTGTTTCTAGCTTGGACTATAGCAGATGGAATTTGGCTTAGAGTCACTCTTCCAAATTT	2187
QY		2161	TTAACTGCGAGCTACTGATGCTGAGACTCTATGCAACAAACATTTAAAGGGTATATTG	2220
Db		2188	TTAACTGCGAGCTACTGATGCTGAGACTCTATGCAACAAACATTTAAAGGGTATATTG	2247
QY		2221	AGAGCTAAATGTTGAAATCTTTTAAATAATGGGAAATTTGAGTTGTTAGTACCAGCTTACAT	2280
Db		2248	AGAGCTAAATGTTGAAATCTTTTAAATAATGGGAAATTTGAGTTGTTAGTACCAGCTTACAT	2307
QY		2281	GATAAAAAACAAACGATTTCTCTGAAGCTCTTGTATATCTCTCGATGAAGCATGTGAAGAC	2340

Db		2308	GATAAAAAACAAACGATTTCTCTGAAGCTCTTGTATATCTCTCGATGAAGCATGTGAAGAC	2367
QY		2341	AAGGCCAATCTTCACTGCTCTCATTTCAGATGAACACAGAGTTATCCCACTTGGAGACAAG	2400
Db		2368	AAGGCCAATCTTCACTGCTCTCATTTCAGATGAACACAGAGTTATCCCACTTGGAGACAAG	2427
QY		2401	GGTTTGTCTTCTCTGCTGAGATTTCTCTCCATTTCCAAAGAGATTTTCTATCTGAATGAA	2460
Db		2428	GGTTTGTCTTCTCTGCTGAGATTTCTCTCCATTTCCAAAGAGATTTTCTATCTGAATGAA	2487
QY		2461	AGAGTTATGTAGCAAAAACAAATTTGAAAAGTGGCACAGGGAATAACAATCCAAATATGTT	2520
Db		2488	AGAGTTATGTAGCAAAAACAAATTTGAAAAGTGGCACAGGGAATAACAATCCAAATATGTT	2547
QY		2521	GACTTGTATGAGGAACAACTCAATGAAGCACTTACTTACCGGAAGCTGTTGATGTT	2580
Db		2548	GACTTGTATGAGGAACAACTCAATGAAGCACTTACTTACCGGAAGCTGTTGATGTT	2607
QY		2581	GATAAATCTATGTTGCTGGAGTAACCAAGATTTACAGCGTCTCTCACGTCTACCTGCTATA	2640
Db		2608	GATAAATCTATGTTGCTGGAGTAACCAAGATTTACAGCGTCTCTCACGTCTACCTGCTATA	2667
QY		2641	CACCTTTATGGAACAACTAGTACACATATAACACAGGCTGCCATTTGTTGGAAGTACAGAT	2700
Db		2668	CACCTTTATGGAACAACTAGTACACATATAACACAGGCTGCCATTTGTTGGAAGTACAGAT	2727
QY		2701	ATTATTTACAGAACTCTGCTGTAATGTTTGGTTCACAGATTTGGATAAGTGGGAAGAAAT	2760
Db		2728	ATTATTTACAGAACTCTGCTGTAATGTTTGGTTCACAGATTTGGATAAGTGGGAAGAAAT	2787
QY		2761	AAAAAATCTGAAAGCACTCTTTGGGCTTGGGAAATATCGGCTCATCAAAATTTGGGGTCTC	2820
Db		2788	AAAAAATCTGAAAGCACTCTTTGGGCTTGGGAAATATCGGCTCATCAAAATTTGGGGTCTC	2847
QY		2821	AATTTGCTACAGGAAGAAAACGTGATTCAGATATATAAATCTTGAACAAACAGTGTGAA	2880
Db		2848	AATTTGCTACAGGAAGAAAACGTGATTCAGATATATAAATCTTGAACAAACAGTGTGAA	2907
QY		2881	GTTCTTTCCATCAGAGGACCTGTGTATATGTACTTGGGCTCATAGCTTAAACCAACAA	2940
Db		2908	GTTCTTTCCATCAGAGGACCTGTGTATATGTACTTGGGCTCATAGCTTAAACCAACAA	2967
QY		2941	GGCTGTGATATTTCTAAAAATGTCAAACTGGGATGCTGTGAGGCATAGTCGCAACATCTG	3000
Db		2968	GGCTGTGATATTTCTAAAAATGTCAAACTGGGATGCTGTGAGGCATAGTCGCAACATCTG	3027
QY		3001	TGGCCAGTGGTTCCAGATGATGTGAAACAACTCTGTATGAACCTTTTCACTATCCCAAGC	3060
Db		3028	TGGCCAGTGGTTCCAGATGATGTGAAACAACTCTGTATGAACCTTTTCACTATCCCAAGC	3087
QY		3061	ACTCTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTGTG	3120
Db		3088	ACTCTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTGTG	3147
QY		3121	CCATCGAGTATGTTTCAATATTTGGAGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT	3180
Db		3148	CCATCGAGTATGTTTCAATATTTGGAGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT	3207
QY		3181	TTCTCTGTATATCAATGAAGATACAGAGCCAACTTTTATGACCCGATCTGGAGCCCAATAAG	3240
Db		3208	TTCTCTGTATATCAATGAAGATACAGAGCCAACTTTTATGACCCGATCTGGAGCCCAATAAG	3267
QY		3241	GATAAAAAATCAATCCCTTTCTTCTTCTAGTAACTTGTGAAGATCTGTATCTTAAAT	3300
Db		3268	GATAAAAAATCAATCCCTTTCTTCTTCTAGTAACTTGTGAAGATCTGTATCTTAAAT	3327
QY		3301	TGCTTACTTTGCTTAAACAAAAACATCGTAGTAGCAGTGTATCCAAAAAGGAGGAAATTA	3360
Db		3328	TGCTTACTTTGCTTAAACAAAAACATCGTAGTAGCAGTGTATCCAAAAAGGAGGAAATTA	3387
QY		3361	TCATCTGAAGTGAAGACAGCAAGGGAATCAGAACACTTACCGAGCCAGTGTGAT	3420
Db		3388	TCATCTGAAGTGAAGACAGCAAGGGAATCAGAACACTTACCGAGCCAGTGTGAT	3447

```
Qy 3421 TTTAATCATAGTGTGATGATTTTACACCCATATCCATCTGTACAGAAACATTAACATTTAGAG 3480
Db 3448 TTTAATCATAGTGTGATGATTTTACACCCATATCCATCTGTACAGAAACATTAACATTTAGAG 3507
Qy 3481 ACTTCATTTATGGGGAATAGACATTTGAAGACACTCTGGTAGTACACCAAGCATTTGGAGAA 3540
Db 3508 ACTTCATTTATGGGGAATAGACATTTGAAGACACTCTGGTAGTACACCAAGCATTTGGAGAA 3567
Qy 3541 AATGACTTAAATTTACCAAGAAATTTTGGTACAGAGAAATACAGAGAAATACCAAGCCGA 3600
Db 3568 AATGACTTAAATTTACCAAGAAATTTTGGTACAGAGAAATACAGAGAAATACCAAGCCGA 3627
Qy 3601 GAGAGGTTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3660
Db 3628 GAGAGGTTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3687
Qy 3661 AATACAGACACTCAACAGTGGGATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3720
Db 3688 AATACAGACACTCAACAGTGGGATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3747
Qy 3721 GTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3780
Db 3748 GTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3807
Qy 3781 ACTAAAACTATTAGACAGCCACTATTGACGCGCACAGTCTAACCATCTGTCTCTCTCC 3840
Db 3808 ACTAAAACTATTAGACAGCCACTATTGACGCGCACAGTCTAACCATCTGTCTCTCTCC 3867
Qy 3841 AATCAAAATTCGGTGTCCCTGGTGTCCCTGGTGTCCCTGGTGTCCCTGGTGTCCCTGGTGTCCCT 3900
Db 3868 AATCAAAATTCGGTGTCCCTGGTGTCCCTGGTGTCCCTGGTGTCCCTGGTGTCCCTGGTGTCCCT 3927
Qy 3901 CAGTCCCTTAAAGCACCCTCTATTGTCTACAAATTAAGTCTAGCAGATTTGACTTTAGT 3960
Db 3928 CAGTCCCTTAAAGCACCCTCTATTGTCTACAAATTAAGTCTAGCAGATTTGACTTTAGT 3987
Qy 3961 TACACAGTTCCTAGAGATGCTTTTGGCTATGCTACACTGAAAGAGACTTACAGCAACAAGA 4020
Db 3988 TACACAGTTCCTAGAGATGCTTTTGGCTATGCTACACTGAAAGAGACTTACAGCAACAAGA 4047
Qy 4021 ATGATCCATCTTATCTCACTGGAAGCTTTGGCTATGCTACACTGAAAGAGACTTACAGCAACA 4080
Db 4048 ATGATCCATCTTATCTCACTGGAAGCTTTGGCTATGCTACACTGAAAGAGACTTACAGCAACA 4107
Qy 4081 ACTGATACCATCACCATGAAAGGCCAACAGTTTGGTCCAGATTTAACACCAAGCAGGTTTC 4140
Db 4108 ACTGATACCATCACCATGAAAGGCCAACAGTTTGGTCCAGATTTAACACCAAGCAGGTTTC 4167
Qy 4141 ATGAAAGCCCTTAAGTTATGATCATTTAGATAAAGAGAGATTTATGAGTCCCTATTATCAAA 4200
Db 4168 ATGAAAGCCCTTAAGTTATGATCATTTAGATAAAGAGAGATTTATGAGTCCCTATTATCAAA 4227
Qy 4201 AATACCCCTGCAAGCATCTTCTCAGTGGCGTCCATGGTGTCCAGTGTCCACATATGGGGGT 4260
Db 4228 AATACCCCTGCAAGCATCTTCTCAGTGGCGTCCATGGTGTCCAGTGTCCACATATGGGGGT 4287
Qy 4261 TCAGATGATTTACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4320
Db 4288 TCAGATGATTTACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4347
Qy 4321 GATATTCCTCTATTTCAGACAAAAACATACACCATCATGATCATGATCGAGGTGCAAGAGCA 4380
Db 4348 GATATTCCTCTATTTCAGACAAAAACATACACCATCATGATCATGATCGAGGTGCAAGAGCA 4407
Qy 4381 TTTGCCCATGATGAGGAGTCTTCCATCTGGAACCTGGAGGTCTTGTAAAAAATCTCTTTT 4440
Db 4408 TTTGCCCATGATGAGGAGTCTTCCATCTGGAACCTGGAGGTCTTGTAAAAAATCTCTTTT 4467
Qy 4441 CACTTGTACGACGAGATGAGTCTTACGGAATATGAAATCAATCCATTCAGATGCC 4500
Db 4468 CACTTGTACGACGAGATGAGTCTTACGGAATATGAAATCAATCCATTCAGATGCC 4527
```

```
Qy 4501 TCTCTGTTTTAGAAAGTACAGAAAGACACTGGACTACAGGAACATACAGATGATAAAGTGC 4560
Db 4528 TCTCTGTTTTAGAAAGTACAGAAAGACACTGGACTACAGGAACATACAGATGATAAAGTGC 4587
Qy 4561 CTTTATTGTTCTGTATTGAAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCATA 4620
Db 4588 CTTTATTGTTCTGTATTGAAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCATA 4647
Qy 4621 TGTAGTCATTTCAGACTTTTCAAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4680
Db 4648 TGTAGTCATTTCAGACTTTTCAAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4707
Qy 4681 AATCCTTTAGAGTGGTTCCTCTAAAGTTTTCGGGATTTTCTGGATCAGTGTGGGGTG 4740
Db 4708 AATCCTTTAGAGTGGTTCCTCTAAAGTTTTCGGGATTTTCTGGATCAGTGTGGGGTG 4767
Qy 4741 TCTCAAGAGGCTCAGCTAGCAGCACCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4800
Db 4768 TCTCAAGAGGCTCAGCTAGCAGCACCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4827
Qy 4801 ATTCCAGATGATACCAATGTCGGTATATCTCTCGCAAGAGAGTTCTTAAGATTTAGTC 4860
Db 4828 ATTCCAGATGATACCAATGTCGGTATATCTCTCGCAAGAGAGTTCTTAAGATTTAGTC 4887
Qy 4861 ATTAAATTTGATAGTGTTCAGTTTCAACTTAAATGTATGATGAGACTGGGCTTTTAAACAATTAAG 4920
Db 4888 ATTAAATTTGATAGTGTTCAGTTTCAACTTAAATGTATGATGAGACTGGGCTTTTAAACAATTAAG 4947
Qy 4921 GAGAAGTATCTCTCAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4980
Db 4948 GAGAAGTATCTCTCAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5007
Qy 4981 TCACACTGCACATTTAGACTTCCGTTGCGGAGTTCATACAGAAATTTTTCAGATGTA 5040
Db 5008 TCACACTGCACATTTAGACTTCCGTTGCGGAGTTCATACAGAAATTTTTCAGATGTA 5067
Qy 5041 CAGTTTCTACAAATGATGAAGAGCAGAGGCTGTGTGGCAACACCAACCAACCAACCT 5100
Db 5068 CAGTTTCTACAAATGATGAAGAGCAGAGGCTGTGTGGCAACACCAACCAACCAACCT 5127
Qy 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127
Db 5128 ATAGTTGATACATCTGCTGAATCCTGA 5154

RESULT 4
AY540053 9297 bp mRNA linear ROD 03-MAR-2004
DEFINITION Mus musculus pianissimo mRNA, complete cds.
ACCESSION AY540053
VERSION AY540053.1 GI:44241266
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 9297)
AUTHORS Shiozaki, C., Woo, J.T., Lindner, J., Shelton, K.D., Zhu, M.Y. and
Magnuson, M.A.
TITLE Mouse pianissimo
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9297)
AUTHORS Shiozaki, C., Woo, J.T., Shelton, K.D. and Magnuson, M.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2004) Molecular Physiology and Biophysics,
Vanderbilt University, 21st and Garland Avenues, Nashville, TN
37232-0615, USA
FEATURES
source
1. .9297
/organism="Mus musculus"
/mol_type="mRNA"
/strain="LAF1"
/db_xref="taxon:10090"
```

CDS	/chromosome="15" 21..5147 /note="similar to Dictyostelium discoideum pianissimo and Drosophila melanogaster CG8002 gene product" /codon_start=1 /product="pianissimo" /protein_id="AA046920.1" /db_xref="GI:44241267" translation="MAAIGRGRSLKNLIRGRNDSEENVPLDLTRPESDNLREILQN VAKLQVSNMKLGHLLNFVKLLCDTGHSEKLGPNVEDIIICLRALLINAEKVRRAA GRLRVLQDSSTLOKVLKYDYLIGKEDIQOSNEVERTQALRLVKMTIVNASL FSSVANSLLAVGNDGLQERDMVRACIALICELALQNEPVVALGGNTILKNVIDC QLSRINEALITITLHLNLHPKTRQYVRADVELERILAPITDHYRHSRDPTEAGQKED REARFLASKMGIIATPFSWAGIINLCKPNSGQISGLIVGICPMIEIRKGLLEVDYDI FRLPLVPVTRDFEIBALLSDPGRFODSWRLSDGFAVEAKTILPHRAISRPLDMNYL ALILSPATRNGLLEGLVEVITNSDDHISVRATILLGELLHMANTILPHSHSHLCLP TLMNAASDIPKEKRLRASAALNLRPHMKRGPKVPYSLHLDHILOKALATHHKR DOYLVRQVDIFVLKDTSEALLINLRSQVLOHKNLDWNLIGITILKWPVNLRNYK DEQLRVRRLLYFYKPSKKLYASLDDLAKSLQVLVGCQTEFLHSEEDGCQYLE DLVDIIVOLMASGVKPERSLQNGLLTISQHYFLFIITGLISLPHGVKMLEKSVF OCLNLCSLKQNDHLKLTVSSLDYSRDGLARVILSKLITAAADCRULTLPHRSLRVLL RANVEFFNNWGIELVQLADNKNTISSEALDIDSEACDKANLHALIQMKPALSHLG DKGILLFLWFSIIPKFSYLNERYGVAKOLEKWHKYNKYVDLIEQOLNEALTYTKR PIDGNVYRSNQRLOPHVYLVPHVQLGVHHTKCHLLEQVSIITELCHNVRTPVL DKWEDIKKLASLWALNGISSNWGLNLOENVIPDILLAKQCEVLISIRGTCCVDL GLIAKTCDDILKCHSWDSVRRSHKHLNVPVDDVEQLCNELSSVPSTLSNSES SRHNSSEAPSMFLEDDRRFGSTSTSTFFLDINEDAEPAFYDRPGPIKDKNSPFF GSSKLKVNRIINSLTITPKHRSDDPKGKLSENKTSNRIRLTLEPSPVDLNHSD FTSSAOKSLOLEPFSVGNKHLEDAGSTPSIGENDLKFPPKSFOTETHRENTSRRLV EGSASSHIKIRSQSFNTDTTSGISMSMSPSRETVAVPTAMDTCGSLSTVSTKT VKTSHYLPQSNHLSLKSNSVILVPPGSHTLIPRRASLKPAPSIATIKSLADCNFSY TSRDAPFATLKLQOQRHPSLISHSEALAPAKOVLFTDITIMKANSFESRLTPSR FMKALUSLDKEDLLSPINNTILQRSRSVRSWSSATYGGSDDYIIGLALPYDINDIF QIKDVPFYQSHVPPDDRRGDPFDHSGAGLGVKNSFHLLRQMSLEIWNIS VHSDASLFLESTEDTGLQEHDDNCLYCVCEILGLOPNSQLSSI CSHSDLDQDIPYSD WCEQITHNPLEVPSKFSIGSCDGSQBEQSASTKSTLELLGKVTIPDDTPWCRI LARKVLRVLVNLSSVSTKCHETGLLTIKKVPQTPDDICTLYSEVSHLLSHCTFRLQ CRFTFQELFQDVQVFLQHEEAEAVLPIPIQPIVDESAES" 2570 /replace="t" 4200 /replace="t"	variation variation
ORIGIN	Query Match 81.1%; Score 4160.4; DB 10; Length 9297; Best Local Similarity 88.5%; Pred. No. 0; Mismatches 586; Indels 6; Gaps 2; Matches 4538; Conservative 0	
Qy	1 ATGCGCGCATCGCGCGCGCGCTCTCTGAAGAACCTTCGAGTACGAGGCGCGAATGAC 60	
Db	21 ATGCGCGCATCGCGCGCGCGCTCTCTGAAGAACCTTCGAAACCTTCGAGGCGCGAATGAC 80	
Qy	61 AGCGGCGAGAGAACGTCGCGCTGATCTGACCCGAGAACCTTCTGATTAATTAAAGAG 120	
Db	81 AGCGGCGAGAGAACGTCGCGCTGATCTGACCCGAGAACCTTCTGATTAATTAAAGAG 140	
Qy	121 ATTCTCAAAATGTGCCAGATTGAGGAGGATCAAAATATGAGAAAGCTAGGCCATCTG 180	
Db	141 ATTCTCAAAATGTGCCAAATTGCAAGGATCAAAATATGAGAAAGCTAGGCCATCTG 200	
Qy	181 AATAACTTTTACTAAGCTTCTTTGGTGTAGATATGGCCACAGTGAAGAAAACTGGCTTTCAC 240	
Db	201 AATAACTTCACTAAGCTTCTTTGGCATATGGCCATAGTGAAGAAAACTGGCTTTAAT 260	
Qy	241 TATGAGATATCATAAATTTGTTGGTGTAGCTTTTATTAAATGAAGCAAAAAGAGTGCGA 300	
Db	261 TATGAAGACATCATAAATTTGTTGGCCTTAGCTTTTATTAAATGAGCAAAAAGAGTGA 320	
Qy	301 GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTTCTCCAGAGGTG 360	
Db	321 GCAGCGGTTGCGAGCGCTGCCTATCTCATCCAAGATTCAGATTTCTCCAGAGGTG 380	

Db 1461 GAGATGAAGAGCGAGGACCTTAAGCCCTACAGCCCTTCATTTAGATCAATCATTCAGAAA 1520
Qy 1501 GCAATGTCACACACACAGAAACGGGATCAGATATCTCCGAGTTTCAGAAAGATATATTTATC 1560
Db 1521 GCAATCGCACTACACACAGCGGAGTCAATATCTTCGAGTTCAAGAAAGATATATTTGTT 1580
Qy 1561 CTTAAGGATACAGAGGAAGCTCTTTAAATTAACCTTAGAGATAGCAGCAAGTCTTCAACAT 1620
Db 1581 CTTAAGGATACAGAGGAAGCTCTTTAAATTAACCTTAGAGATAGCAGCAAGTCTTCAAGCAT 1640
Qy 1621 AAAGAGAACTTGAATGGAATCTTATAGGACCATTTCTTAAGTGGCCAAATGTA 1680
Db 1641 AAAGAGAACTTGAATGGAATCTTATAGGACCATTTCTTAAGTGGCCAAATGTA 1700
Qy 1681 AATCTAAGAACTATAAAGATGAACAGTTTACACAGGTTTGTACGAAGACTACTTTATTTT 1740
Db 1701 AATCTAAGAACTATAAAGATGAGCAGTTGCACAGGTTTGTGCGCAGACTCTTTACTTT 1760
Qy 1741 TACAAGCCAGCAGTAATATATATGCCAACCTGGATCTGGATTTTGGCAAGGCCAAACAG 1800
Db 1761 TACAAGCCAGCAGCAAACTGTACGTAGTCTGGATCTGGATCTGGCAAGGCCAAACAG 1820
Qy 1801 CTCAGGTTGTAGTGGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGATGGGCAA 1860
Db 1821 CTCAGGTTGTAGTGGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGATGGGCAA 1880
Qy 1861 GGCTACTTAGAAGATCTAGTAAGAGATATTTGTCAGTGGCTCAATGCTTCAATCTGGAAATG 1920
Db 1881 GGATACTTAGAAGATCTCGTGAAGATATTTGTCAGTGGCTCAATGCTTCAATCTGGAGTG 1940
Qy 1921 AAACCCGAAAGAGCTTTCAAAATAAGTTTATTTGACCACCCCTTAGTCAACACTACTTTT 1980
Db 1941 AAACCTGAGAAAGCCTTCAGAAATATGGCTTACTGACCAACCTTAGTCAACACTACTTTT 2000
Qy 1981 TTATTTATGGAACACTTTCTTGGCCACCTCATGAGTTTAAATGCTGGAAAAATGCAATG 2040
Db 2001 TTATTTATGGAACACTTTCTTGGCCACCTCATGAGTTTAAATGCTGGAAAAATGCAATG 2060
Qy 2041 GTATTTAGTGTCTCTTAATCTTTGCTCTTGAATAACCAAGATCACTTCTTAAACCTT 2100
Db 2061 GTCTTTAGTGTCTCTTAATCTTTGCTCTTGAATAACCAAGATCACTTCTTAAACCTT 2120
Qy 2101 ACTGTTCTAGCTTGGACTATAGCAGAGATGGATGGCTAGAGTCACTCTTTCCAAATTT 2160
Db 2121 ACTGTTCTAGCTTGGACTATAGCAGAGATGGATGGCTAGAGTCACTCTTTCCAAATTT 2180
Qy 2161 TTAAGTCAAGCTAGAGTCTGAGACTATGCAACAAACATTTAAGGGTATTTATG 2220
Db 2181 CTCAGGCGAGCCAGAGATGCTTGAGACTGTATGCAACAAACATTTAAGGGTATTTATG 2240
Qy 2221 AGAGCTAATGTTGAATCTTTAAATTAATGGGGAATTTGATGTTAGTGACCCAGCTACAT 2280
Db 2241 AGAGCTAATGTTGAATCTTTAAATTAATGGGGAATTTGATGTTAGTGTTACTGACTACAT 2300
Qy 2281 GATAAAACAAACAGATTTCTCTGAAGCTCTTGATATCTCGATGAGCAGTGTGAAGAC 2340
Db 2301 GATAAAACAAACAAATTTCTCTGAAGCTCTTGATATTTCTTGATGAAGCTTTGGAAGAC 2360
Qy 2341 AAGGCCAATCTTGATGCTCTCATTCAGATGAACAGGGTTATCCGACCTTGGAGACAAG 2400
Db 2361 AAGGCCAATCTTGATGCTCTCATTCAGATGAAGCCGCTTATCCGACCTTGGAGACAAG 2420
Qy 2401 GGTGTTGCTCTCTGCTGAGATTTCTCTCAATTCGAAAGGATTTCTCTATCTCAATGAA 2460
Db 2421 GGTGTTGCTCTCTGCTGAGATTTCTCTCAATTCGAAAGGATTTCTCTATCTCAATGAA 2480
Qy 2461 AGAGGTTATGTAGCAAAACAAATGGAAAGTGGCAGAGGAAATCAACCTCCAAATATGTT 2520
Db 2481 AGGGGTTATGTAGCAAAACAAATGGAAAGTGGCAGAGGAAATCAACCTCCAAATATGTT 2540
Qy 2521 GACTTGTATGAGGAACAACTCAATGAAGCACTTACTTACCGGAAGCCCTGTTGATGGT 2580

Db 2541 GACTTGTATGAGGAACAACTTAAATGAAGCCCTCACTACATACAGGAAGCCCTATTGATGTT 2600
Qy 2581 GATAACTATGTTGTCGAGTAAACCAAGATTTACAGCGTCTCTCAGTCTTACCTGCTATTA 2640
Db 2601 GATAACTATGTTGTCGAGTAAACCAAGATTTACAGCGTCTCTCAGTCTTACCTGCTATTA 2660
Qy 2641 CACTTTATGACAACTAGTACACATATAAAACAGGCTGCCATTTGTTGGAAAGTACAGAAAT 2700
Db 2661 CACTTTATGACAACTAGTACACATATAAAACAGGCTGCCATTTGTTGGAAAGTACAGAAAT 2720
Qy 2701 ATTTATACAGAACTCTGTCGTAATGTTGCTACACAGATTTGGATTAAGTGGGAAGAAAT 2760
Db 2721 ATCAATACAGAGCTCTGTCAATGTTGCGACTCCAGACTTGGCAAAATGGGAAGCAAT 2780
Qy 2761 AAAAACTGAAGCATCTCTTTGGGCCCTTGGGAAATATCGGCTCATCAAAATTTGGGGCTCTC 2820
Db 2781 AAGAAGCTGAAGCCTCTCTTTGGGCCCTTGGGAAATATTTGGCTCATCGAATTTGGGGCTT 2840
Qy 2821 AATTTGCTACAGGAAGAAACGCTGATTCAGATATATCTTAAACCTTGAACAAACAGTGTGAA 2880
Db 2841 AATTTGCTACAGGAAGAAATGATCCAGATATATCTTAAACCTTGAACAAACAGTGTGAG 2900
Qy 2881 GTTCTTTCCATCAGAGGACCTGTGTATATCTTGGGCTCATAGCTTAAACCCAAACAA 2940
Db 2901 GTCTTTCCATCAGAGGACCTGTGTATATTTCTTGGGCTCATAGCTTAAACCCAAACAA 2960
Qy 2941 GGCTGTGATATTTAAATATGTACAACTGGGATGCTGTGAGGCAATAGTCGCAAAACATCTG 3000
Db 2961 GGCTGTGATATTTGAATATGTACAGCTGGGATTTCTGTAAAGCAGTCCGCAAGCATCTA 3020
Qy 3001 TGGCCAGTGGTTCCAGATGATGTGGAAACAACTCTGTAAATGAACCTTTCATCTATCCCAAGC 3060
Db 3021 TGGCCAGTGGTTCCAGATGATGTGGAGCAGCTCTGCAATGAGCTCTCATCTGTCCCAAGC 3080
Qy 3061 ACTCTAAGTTTGAACCTGGAGTCAACAGCTCTTAGACATATAATAGTGAAGTGAATCTGTG 3120
Db 3081 ACCCTCAGCTTGAACCTCTGAGTCGACCACTCTAGACATATAACAGTGAAGTGAATCTGCA 3140
Qy 3121 CCATCGAGTATGTTTCAATTTGGAGGATGACCGGTTTGGCAGCAGCTCTCTACTAGTACATTT 3180
Db 3141 CCATCAAGCATGTTTCAATCTGAGGAGCAGCCGCTTGGCAGCAGCTCTCAACAGACATTT 3200
Qy 3181 TTCCTTGATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCCAATAAG 3240
Db 3201 TTCCTTGATATCAATGAAGATGCTGAGCCAGCACTTTTATGACCGACCTCTGGACCCATAAA 3260
Qy 3241 GATAAAAATTCATCTCCCTTTCTTCTTCTAGTAAACCTTGTGAAGATCTGTATCTTAAAT 3300
Db 3261 GATAAAAATTCATCTCCCTTTCTTCTTCTGCTCCAGCAACCTTGTAAAGATCTGTATCTTAAAT 3320
Qy 3301 TCGCTTACTTTGCTTAAACAAACATCTGTAGTACAGTATCCAAAGGAGGGAATTA 3360
Db 3321 TCATTTACTTTGCTTAAACAAACCTGAGTACAGTATCCAAAGGAGGGAATTA 3380
Qy 3361 TCATCTGAAGTAAAGCAAGCAACAGCGAATCAGAACTTACGAGCCAGTGTGAT 3420
Db 3381 TCATCTGAAGTAAAGCAAGCAACAGCGAATCAGAACTTACGAGCCAGTGTGAT 3440
Qy 3421 TTTAATCATAGTATGATTTTACCCATATCTCTGACAGAAACATTAATTAATAG 3480
Db 3441 TTGAATCAGAGCAAGACTTTCACA---AGCTCTCAGCAGCAAGAAATCATTAACAGTAGAG 3497
Qy 3481 ACTTCAATTTTGGGAAATAGCATTGAAGACACTGTAGTACACCAAGCATTTGGAGAA 3540
Db 3498 CTTTCTCTTGGGAAATAGCATTGAAGACCGGGCAGCACTCCAAGTATTTGGAGAG 3557
Qy 3541 AATGACTTAAAAATTCACCAAGAAATTTTGGTACAGAAATCACAGAGAAATATCAAGCCGA 3600
Db 3558 AATGACTTAAAAATTTCCCAAGAGTTTGGTACCGAGACTCACAGAGAAATATCAAGCCGA 3617
Qy 3601 GAGAGGTTAGTATGAGAAAGTTCAACGAGCTCAATATGAAGATACGTAGCCAAAGTTTC 3660
Db 3618 GAAAGTTGTTGTTGGAAGGCTCAGCAAGCTCCCAATCAAGATCCGAGCCAGAGCTTT 3677

QY 3661 AATACAGACACTACAAAGTGGCATAAGTTCAATGAGCTCAAGTCTCTTTCAGGAGACA 3720
Db 3678 AACACGACACTACGACGAGTGGCATCAGCTCCATGAGCTCGAGCCCTCCCGAGAGACA 3737
QY 3721 GTAGTGTAGATGCTCAACTATGGAACAGAGACTGTGGAAGCATGAGTACTGTGTAAGT 3780
Db 3738 GTGCTGTGCGACCCACAGACTATGGAACAGAGACTGTGGAAGCTGAGTACCGTGGTAAAGT 3797
QY 3781 ACTAAACATTAAGACAGACCACTATTTGACGCGCACAGCTCTAACCATCTGTCTCTCC 3840
Db 3798 ACTAAACATTTAAGACGAGCCACTATTTGACACCAAGCTCTAACCATCTGTCTCTCC 3857
QY 3841 AAATCAAAATTCGCTGTCCCTGGTCCCTCCAGGTTCTTCTCATACGCTTCCCTAGAAGACA 3900
Db 3858 AAATCGAATCTAGTGTCCCTGGTCCCAACAGGTTCTTCTCACCCCTTCTTAGAGAGCC 3917
QY 3901 CAGTCCCTTAAAGCACCCCTCTATTTGCTACAAATTAAGGCTTAGCAGATTTGATTTAGT 3960
Db 3918 CAGTCCCTTAAAGCGCCCTCCATTGCAACAATAAAGGCTTAGCAGATTTGCACTTCAGC 3977
QY 3961 TACAAAGTCTAGAGATGCTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAGA 4020
Db 3978 TACAAAGCTCTAGAGATGCTTTGGCTATGCTACACTGAAAAGGTTTACAGCAGCAAGA 4037
QY 4021 ATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTT 4080
Db 4038 ATGCACCGCTCTTGTCTCTACTCGGAAGCTTTGGCGTCCGCGCAAAAGATGTCTTTTT 4097
QY 4081 ACTGATACCATCACTCAAGAGCCAAAGTCTTGTAGTCCAGATTAACACCAAGCAGTTTC 4140
Db 4098 ACCGACACCATCACTCAAGAGCCAAAGCTTTGAAATCCAGACTGACGCAAGCAGTTT 4157
QY 4141 ATGAAAGCCTTAAGTATGCAATCTAGATAAAGAGATTTATGAGTCTTATTAATCAA 4200
Db 4158 ATGAAAGCTTTAAGTATGCGTCAATAGATAAAGAGATTTACTAAGTCTTATTAATCAC 4217
QY 4201 AATACCTCGAACGATCTTCTCAGTGGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCT 4260
Db 4218 AATACCTCGAGCGCTCTCTCAGTGGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCT 4277
QY 4261 TCAGATGATTAATGATGCTGCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4320
Db 4278 TCCGATGATTAATGATGCTGCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4337
QY 4321 GATATTTCCCTATTTTTCAGACAAACATACACCACTGATGATGATGATGATGATGATGATG 4380
Db 4338 GATGTTTCCCTATTTTCAATCAAAAGCAGTGGCTTCCACCATGATGATGATGATGATGATG 4397
QY 4381 TTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4440
Db 4398 TTTTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4457
QY 4441 CACTTGTACGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4500
Db 4458 CACTTGTGCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4517
QY 4501 TCTCTGTTTTAGAAAGTACAGAGACACTGGCTACAGAAATACAGAAATACAGAAATACAG 4560
Db 4518 TCCCTGTCTTAGAGACACAGAAAGACTGGGCTGCGAGAGACACAGGATGACAAATGCG 4577
QY 4561 CTTTATTTGCTGTGATTTGAAATCTGGGTTTCCAGCCACAGCAACCACTGAGTGAATA 4620
Db 4578 CTTTATTTGCTGTGATTTGAAATCTGGGTTTCCAGCCACAGCAACCACTGAGTGAATA 4637
QY 4621 TGTAGTCACTTCAAGATATTTCCATATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 4680
Db 4638 TGCAGTCACTCCGATCTTCAAGATATTTCCGATATTTCCGATATTTCCGATATTTCCGAT 4697
QY 4681 AATCCTTTAGAGTGGTTCCTCTTATGTTTTCCGAGATTTTCTGATGATGATGATGATGATG 4740
Db 4698 AACCCCTTAGAAGTGGTTCCTCTTATGTTTTCCGAGATTTTCTGCTGATGATGATGATG 4757

QY 4741 TTCT---AAGAAGCTCAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4797
Db 4758 TCTCAGGAAGAGGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4817
QY 4798 ACATTTCCAGATGATACACCATGTCGCGTATCTCTTCCGAAAGAGTTCCTAGATTA 4857
Db 4818 ACCATTCCAGATGATACACCATGTCGCGGATCTCTTCCGAAAGAGTTCCTAGATTA 4877
QY 4858 GTCATTAAATTTGAGTAGTTTCAATTAATGTCATGAGCTGGGCTTTTAAACAAT 4917
Db 4878 GTTGTCAATTTGAGTAGTTTCAATTAATGTCATGAGCTGGGCTTTTAAACAAT 4937
QY 4918 AAGGAGAGTATCTTCAAAACATTTGATGATGATGATGATGATGATGATGATGATGATG 4977
Db 4938 AAGGAGAGTATCTTCAAAACATTTGATGATGATGATGATGATGATGATGATGATGATG 4997
QY 4978 CTGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5037
Db 4998 TTGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5057
QY 5038 GTCACAGTTTCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5097
Db 5058 GTCACAGTTTCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5117
QY 5098 CCTATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5177
Db 5118 CCTATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5197

RESULT 5
AY497009
LOCUS Mus musculus AVO3 mRNA, complete cds. 5127 bp mRNA linear ROD 13-JAN-2004
DEFINITION
ACCESSION AY497009
VERSION AY497009.1 GI:40737169
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE 1 (bases 1 to 5127)
AUTHORS Loewith, R.J. and Hall, M.N.
TITLE Mouse mAVO3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5127)
AUTHORS Loewith, R.J. and Hall, M.N.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2003) Biochemie, University of Basel Biozentrum,
Klingelbergstrasse 50/70, Basel, BS 4056, Switzerland
FEATURES
source
1..5127
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H"
/db_xref="taxon:10090"
/chromosome="15"
/cell_line="C2C12"
1..5127
/note="mAVO3; binds mTOR; similar to yeast AVO3 and Dictyostelium discoideum pianissimo"
/codon_start=1
/product="AVO3"
/protein_id="AA889074.1"
/db_xref="GI:40737170"
/translation="NAATIGRSLNLRVRGRNDSEENVPLDTRPSDNLREILQN
VAKLQVSNMRKLGHLNFTKLLCDIGSEKLGNYEDIIICLRLALINEAKEVRAA
GURKLYLIQDSILQKLVKLVKLIARCIDIQSNVEVQALRLVREKMTVNSL
PFSSVANSIYVNDGLQERDMVQACIICELALQNEVVALRGMLNLIKVNIDC
QLSRNEALITITLHLNPKTRQYRADVELERLAPYDFYRHSPTAGQLKED
REARFLASKMGIATFRSWAGIINCKPNSGSIQSLIGVLCIPNMEIRGLLEVDI
FLPLPVTDFEALLSVDPGRFODSWRLSDGFVAAEAKTILPHRARSPLMDNYL
ALILSAFTENGIEGLVEVITNSDDHISVRATILIGELHLMANTILPHSHHLCPL
TLNMAASFDPKEXELRASAALNCLNRPHEMKRGPFPKPSLHLDHIIQKATATHKR

CDS
1..5127
/note="mAVO3; binds mTOR; similar to yeast AVO3 and Dictyostelium discoideum pianissimo"

DOYLRVQKDI FVLKDTBEALLINLRDSQVLQHKENLDWDMNLIIGTILKWPNNVLRNKK
DEQLHRFVRLILYFKPSSKLYASLDLDDAKSQLTVVGCQFTBFLFSEEDGCGYLE
LDXVDI VQMLNASSGVKPERSLQNNGLTTLISQHYFLPFGTSLPHGVKMLKRCVSF
QCLNMLCKNQDHLILVTLVSSLDYSRDLARLILTAATAADACRLAYATKHLRVLL
RANVEFFNNWGIELLVTLQHDKNKTIISSEALDILDEACEDKANHALIOMKPLSHLG
DKGLLLARPLSI PKGFSYLNKRGVVAQOLEKMKHEVNSKYVDLIEBQNEALITYRK
PIDGNNYRRNQRLRPHVYLPVHLVQVLVHKHTGCHLLEVSQSIITELCHNVRTPD
DKWEDI KKLKASLWALNGISSNMLNLLQBEENVI PDILKLAQCEVLISRGTCVYVL
GLIAKTGCGDILKCHSDVSRHKHLPVVPDDVEQLCNELSSVSTSLNSESST
SRHNSSESAPMFLEDDRFSGTSTSTPFLDINEADPAFYPDRPGPKDKNSFPFF
GSSKLKVNRI LNSLTLPTKHKRSSDPKGGKLSSENKTSNRRIRTLTLEPSVDLHNSD
FTSSAOKSLILEPSFVGNKHLEDAGTPSIGENDLKPFPKSPFTEHRENTSRRLV
EGSSASHIKIRSQFNIDTTSGISMSSSPSRETVAVDPTAMDTCGSLSTVVSXTK
TTSHTYLPQSNHLSLKSNSVSLVFPSSHTLPRRAQSLKAPSIATIKSLADCNFY
VSSDRTGYATLKLQORHPPLSHSEALAPAKDVLFTDTITWKNANSFESRLTPSR
FMKALSVALDKBLLQSP INHNTLQRSSVSRMSSVATYGGSDDYI GLALPVDINDPI
IKDVPYQOSKHPPDDRGARFMSHDGAGLSGAGLVKNSPHLLQOMSLTEIMS
VHSDASLPLESTEDTGLQEHDDNCLYCVCIELGFPQSNQLSSI CHSDLOQDIPSD
WCBQTIHNPVPSKSGISGSDGASQEGSASSTKSTELLGLGVKTIIPDTPMCR
LLRKEVLRLVNLSSVSTKCHETGLLTIKEKYPOTFDDICLYSEVSHLLSHCTFRLQ
CRPFIQLFQDQVQLQMHAEAVLAIPPIQPIVDESABE"

ORIGIN

Query Match 81.1%; Score 4158.8; DB 10; Length 5127;

Best Local Similarity 88.4%; Pred. No. 0;

Matches 4537; Conservative 0; Mismatches 587; Indels 6; Gaps 2;

Qy	1	ATGGCGCGATCGGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC	60
Db	1	ATGGCGCGATCGGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC	60
Qy	61	AGCGGAGGAGAACGTCGCCCTCGATCTGACCCGAGAACCTCTCGATAACTTAAGAGAG	120
Db	61	AGCGGAGGAGAACGTCGCCCTCGATCTGACCCGAGAACCTCTCGATAACTTAAGAGAG	120
Qy	121	ATTCTCCAAATGTGGCCAGATTGCGGAGATATCAATATGAGAACCTAGGCGCATCTG	180
Db	121	ATTCTCCAAATGTGGCCAAATTCGAGGAGTATCAATATGAGAACCTAGGCGCATCTG	180
Qy	181	AATAACTTTACTAAGCTCTTTGTGATATGCGCACAGTGAAGAAAACCTGGGCTTTAC	240
Db	181	ATAACTTTACTAAGCTCTTTGTGATATGCGCATATGCGCAATGAGAAAACCTGGGCTTTAAT	240
Qy	241	TATGAGGATATCATATTTGTTGCGGTAGCTTTTAAATGAAGCAAAAGAGTGCGA	300
Db	241	TATGAGGATATCATATTTGTTGCGCTTAGCTTTTAAATGAGCAAAAGAGTGAAGA	300
Qy	301	GCAGCAGGCTACGAGCGCTTCGATATCTATCCAGACTCCAGTATTTCTCCAGAGGTG	360
Db	301	GCAGCAGGCTTCGAGCGCTTCGCTATCTATCCAGAGTTCAGTATTTCTCCAGAGGTG	360
Qy	361	CTAAATTTGAAGTGGACTATTTAATAGTGGTGCATTTGACATACACAGAGCAACGAG	420
Db	361	CTAAACTGAAAGTGGACTATTTAATAGTGGTGCATTTGACATTCAGAGCAACGAG	420
Qy	421	GTAGAGGAGCACAAAGCACTTCGATTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG	480
Db	421	GTGAGAGGACACAAAGCCCTTCGCTTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG	480
Qy	481	TTTCTAGTTCTGTGACCACTCATATTAATTCGAGTTGGAATGATGAGCTTTCAAGAAGA	540
Db	481	TTTCTAGTTCTGTGGGCAACTCATATTAATTTGCGGTGTGAATGATGAGCTTTCAAGACGA	540
Qy	541	GACAGATGTTCGAGGAGTGCATTTGATTCGACTAGCACTAGCACTTCAGATCCAGAG	600
Db	541	GACAGATGTTCGAGGCTTGATCGCCATCATCTGTGAACCTAGCACTTCAGAACCCAGAG	600
Qy	601	GTGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAAATGTGATTGATTGCAATTA	660
Db	601	GTGTAGCCCTTCGAGGAGGACTAAACACCATACTGAAAATGTGATTGATTGCAATTA	660
Qy	661	AGTCAATAATAGGCGCTTAATTTACTACATTTTGCACCTTCTTAATCATCCAAAGACT	720
Db			

Db	661	AGTCAATAATAGGCGCTCTCATTACTACAAATTTTGCACCTTCTTTAATCATCAAGACC	720
Qy	721	CGGAGTATGTCGAGCTGATGTAGAAATTTAGAGAGAAATTTTAGCACCTTACTATGATTTT	780
Db	721	CGGAGTATGTCGAGCTGATGTAGAAATTTAGAGAGAAATTTTAGCACCTTACTATGATTTT	780
Qy	781	CACATACAGACATGTCGAGATACAGCTGAAAGAGCAGCTCAAAGAGACAGAGAACGACA	840
Db	781	CACATACAGACATGTCGAGATACAGCTGAAAGAGCAGCTCAAAGAGACAGAGAACGACA	840
Qy	841	TTTCTAGCCAGTAAATTTGGGAAATCATAGCAACATTCGATCATGGCAGGTATTATTAAT	900
Db	841	TTTCTAGCCAGTAAATTTGGGAAATCATAGCAACATTCGATCATGGCAGGTATTATTAAT	900
Qy	901	TTATGTAACCTGGAATTTCTGGGATCCAGTCTCTAATAGGAGTACTTTTGATACCAAT	960
Db	901	TTATGTAACCTGGAATTTCTGGGATACAGTCTCTAATTTGGAGTACTTTTGATACCAAT	960
Qy	961	ATGAAAATAAGGCGAGGTCTACTTTGAAAGTCTTTATGATATATTTTCTCTCTACCT	1020
Db	961	ATGAAAATAAGGCGAGGTCTCTCGAAGTCTCTATGACATCTTTCTCTCTCTCTCTCTCT	1020
Qy	1021	GTGTGATGAGGAGTTCATAGAGCACTACTCAGTGTAGATCCAGGAGGTTCACAGAC	1080
Db	1021	GTGTGATGAGGAGTTCATAGAGCACTACTCAGTGTAGATCCAGGAGGTTCACAGAC	1080
Qy	1081	AGTTGAGGCTTTTCAGATGCTTTGTGCGAGCTCAGGCAAAATCTATTCTTCTCTCATCT	1140
Db	1081	AGTTGAGGCTTTTCGATGCTTTGTGCGAGCTCAGGCAAAATCTATTCTTCTCTCATCT	1140
Qy	1141	GCCAGATCCAGGCGAGACCTCATGGATAATTTATTTGGCAGTACTCTCTGCAATTTAT	1200
Db	1141	GCTAGATCCAGGCGAGACCTCATGGACAAATTTATTTGGCTCTGTACTCTCTGCAATTTAT	1200
Qy	1201	CGTAATGAGCTTTTAGAGGCTCTAGTTGAGTGTATTAACAAACAGTGTATCTCTCA	1260
Db	1201	CGTAATGAGCTTTTAGAGGCTTTGGTCAGGTGTATTAACAAACAGTGTATCTCTCA	1260
Qy	1261	GTAGAGCTACCACTCTTTTAGGAGAGCTTTACATATGCGCAACACAAATTTCTCTCAT	1320
Db	1261	GTAGGCGGACCATCTCTTTTAGGAGAGCTCTGTGATATGGCAACACCATCTCTGCTCAT	1320
Qy	1321	TCACATAGCCATCATTTTACATCTGCTTCCCAACCTTAATGAATATGGCTGCACTCTTGAT	1380
Db	1321	TGCGATAGCCACCATCTACACTGCTTCCCAACCTCTCATGAACATGGCTGCTCTTGAT	1380
Qy	1381	ATCCCAAGAAAAGAGACTGCGAGCCAGTGCAGCTTGAACCTGTTTAAACGCTTCCAT	1440
Db	1381	ATTTCCCAAGAAAAGAGGCTGCGAGCCAGTGCAGCCCTGAACTGTTTGAACGCTTCCAT	1440
Qy	1441	GAAATGAAGAAAAGAGGAGCTTAAGCCCTTATAGTCTTCTTATAGACCATTTATTCAGAAA	1500
Db	1441	GAGATGAAGAAAAGAGGAGCTTAAGCCCTTATAGTCTTCTTATAGATCATTTATTCAGAAA	1500
Qy	1501	GCAATTTGCAACACACCAAGAACCGGATCAGTATCTCCGAGTTTCAGAAAGATATATTATC	1560
Db	1501	GCAATTCGCAACTCACCAAGCGGATCAGTATCTTCGAGTTTCAGAAAGATATATTATC	1560
Qy	1561	CTTAAGGATACAGAGGAGCTCTTTTAATTAACCTTAGAGATAGCCAGCTCTTCAACAT	1620
Db	1561	CTTAAGGATACAGAGGAGCTCTTTTAATTAACCTTAGAGATAGCCAGCTCTTCAACAT	1620
Qy	1621	AAAGAGATCTGAAATTTGGAATCTTATAGGAGCACTTCTTAAGTGGCCAAATGTA	1680
Db	1621	AAAGAGATCTTGAATTTGGAATCTGATTTGGAGCACTTCTTAAGTGGCCAAATGTA	1680
Qy	1681	AATCTAAGAAACTATAAAGATGAACAGTTACACAGGTTTGTGCAAGACTACTTTATTTT	1740
Db	1681	AATCTAAGAAACTATAAAGATGAACAGTTACACAGGTTTGTGCAAGACTACTTTATTTT	1740
Qy	1741	TACAGCCCGAGGAGTAAATTTATGCAACTGTCGATCTGGAATTTTGCAGAGGCGCAACAG	1800
Db	1741	TACAGCCCGAGGAGTAAATTTATGCTAGTCTGATCTGGAATTTTGCAGAGTCTTCAAGCAG	1800

QY	1801	CTCAGCGTTCTAGTTCGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGSCAA	1860
Db	1801	CTCAGAGTTGGTGTGCTAGTTTACAGAAATTTCTGCTCGAGTCTGAAGAGATGGSCAA	1860
QY	1861	GGCTACTAGGAATCTAGTAAAGGATATTGTTTCAGTGGCTCAATGCTTCATCTGGAATG	1920
Db	1861	GGATACTAGGAATCTCGTGAAGATATTGTTTCAGTGGCTCAATGCTTCATCTGGAGTG	1920
QY	1921	AAACCCGAAGAAGTCTTCAAAATATAGTTTATTTGACCAACCTTCTAGTCAACACTATTTT	1980
Db	1921	AAACCTTGAGAAAGCCCTTCAGAAATAATGGCTTCTGACCAACCTTCTAGTCAACACTATTTT	1980
QY	1981	TTATTTATTTGGAACACTCTTCTGCAACCTCATGGAGTTAAATGCTGGAAAAATCCAGT	2040
Db	1981	TTATTTATTCGAAACACTCTCTTGTCACTCTCATGGAGTCAAAATGCTGGAAAAATCCAGT	2040
QY	2041	GTATTTTCAGTGTCTCCTTAATCTTTGCTCCTTGAATAAACCAAGATCACTTGTCTAAACCTT	2100
Db	2041	GTCTTTTCAGTGTCTCCTTAATCTTTGCTCCTTGAATAAACCAAGATCACTTGTATAAACTC	2100
QY	2101	ACTGTTTCTAGCTTGGACTATACAGAGATGGATTTGGCTAGATCATCTTTCCAAAATTT	2160
Db	2101	ACTGTTTCCAGCCTAGACTATACAGAGATGGATTTAGCCAGAGTCAATCTTTCAAAAATC	2160
QY	2161	TTAACTGCAGTACTGATGCTGCAGACTCTATGCAACAAACATTTAAAGGGTATTATTG	2220
Db	2161	CTCACGGCAGCCACAGATGCTTGCAACTGTATGCAACAAACATTTAAAGGGTATTATTG	2220
QY	2221	AGAGCTAAATGTTGAAATCTTTTAATAATTTGGGAATTTGAGTTGTTAGTGACCCAGCTACAT	2280
Db	2221	AGAGCTAAACGTTGAAATCTTTCAACAATTTGGGAATTTGAGTTTACTAGTAACTCAGCTACAT	2280
QY	2281	GATAAAAACAAACGATTTCTCTGAAGCTCTTGATATCCTCGATGAAGATGTGAAGAC	2340
Db	2281	GATAAAAACAAACATTTCTCTGAAGCGCTGGATATTTCTTGATGAAGCTGTGAAGAC	2340
QY	2341	AAGGCCAATCTTCACTGCTCAATGAAGCACTTACTTACCGGAAGCCTGTTGATGTT	2400
Db	2341	AAGGCCAAATCTTCACTGCTCAATGAAGCACTTACTTACCGGAAGCCTGTTGATGTT	2400
QY	2401	GGTTTGCTTCTCTGCTGAGATTTCTCTCAATTTCCAAAAGGATTTTCTTCTGTAATGAA	2460
Db	2401	GGCTTGCTTCTCTGCTGAGATTTCTTCCATTTCCGAAAGGATTTTCTTCTTACCTGTAATGAA	2460
QY	2461	AGAGGTTATGTAGCAAAACAAATTTGGAAGAGTGGCAAGGGAATACAACTCCAAATATGTT	2520
Db	2461	AGGGGTTATGTTGCAAAACAGTTTGGAAAAAGTGGCAAGGGAATATTAATTCAAAGTATGTT	2520
QY	2521	GACTTGATTTGAGGAACAACCTCAATGAAGCACTTACTTACCGGAAGCCTGTTGATGTT	2580
Db	2521	GACTTGATTTGAGGAACAACCTTATGAAGCCCTCACTACATACAGGAAGCCTTATGATGTT	2580
QY	2581	GATAACTATGTTTCGTCCGAGTAAACCAAGATTAACAGCGTCTCTCACTGCTTACCTGCTATA	2640
Db	2581	GATAACTATGTTTCGTCCGAGTAAACCAAGATTAACAGCACTACAGCGCCTCATGCTTACCTGCCGTA	2640
QY	2641	CACCTTTATGGCAACTAGTACACCATAAACAGGCTGCCNTTGTGGAAGTACAGAAAT	2700
Db	2641	CACCTTTATGGCAACTAGTGTACACCAACAAACAGGCTGCCNTTACTGGAAGTACAGAGT	2700
QY	2701	ATTATTTACAGAACTCTGTGCTAATGTTCTGTACACCAAGATTTGGATTAAGTGGGAAGAAAT	2760
Db	2701	ATCATACAGAGCTCTGTCAATGTTTCGCACTCCAGACTTGGACAAATGGGAAGACATTT	2760
QY	2761	AAAAAATGAAAGCATCTCTTTGGGCTTTGGGAAATATCGGCTCATCAAAATGGGGTCTC	2820
Db	2761	AAGAAGCTGAAAGCCCTCTCTTTGGGCTTTGGGAAATATTTGGCTCATCTGAAATGGGGCTT	2820
QY	2821	AATTTGCTACAGGAAGAAACGTTATCCAGATATATACTAAACCTTGCAAAAACAGTGTGAA	2880
Db	2821	AATTTGCTACAGGAAGAAATGTGATCCAGATATATACTAAACCTTGCAAAAACAGTGTGAG	2880

QY	2881	GTTCCTTTCCATCAGAGGGACCTGTGTATATGTACTTTGGCTCATAGCTAAAAACCAACAA	2940
Db	2881	GTTCCTTTCCATCCGAGGGACCTGTGTATATGTACTTTGGCTCATAGCTAAAAACCAACAA	2940
QY	2941	GGCTGTGATATTCTTAAATGTCAACACTGGGATGCTGTGGGCATAGTGCACAACTCTG	3000
Db	2941	GGCTGTGATATTCTTAAATGTCAACAGCTGGGATTTCTGTAAAGCACAGTGCGAAGCATCTA	3000
QY	3001	TGGCCAGTGGTTCCAGATGATGTGGAAACAACTCTGTGTAATGAACCTTTCATCTATCCCAAGC	3060
Db	3001	TGGCCAGTGGTTCCAGATGATGTGGAGCAGCTCTGCAATGAGCTCTCATCTGTGCCAAGC	3060
QY	3061	ACTCTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTGTG	3120
Db	3061	ACCTCAGCTTGAACCTCTGAGTCGACCACTCTAGACATAACAGTGAAGTGAATCTGTGCA	3120
QY	3121	CCATCCAGTATGTTTCATATTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT	3180
Db	3121	CCATCAAGCATGTTTCATGCTGGAGGACGACCGCTTCGGCAGCACCTCCACAGCACATTT	3180
QY	3181	TTCCCTTGATATCAATGAAGATACAGAGCCAAATTTTATGACCGATCTGGACCCATAAG	3240
Db	3181	TTCCCTTGATATCAATGAAGATGCTGAGCCAGCATTTTATGACCGACCTCGACCCATAAA	3240
QY	3241	GATAAAAATTCATTCCTTTCTTGTCTTAGTAAACTTTGTGAAGATCGTATCTTTAAAT	3300
Db	3241	GATAAAAATTCATTTCCCTTTTGGCTCCAGCAAACTTTGTAAAGNATCGTATCTTAAAT	3300
QY	3301	TGCTTTAATTTGCTTAAACAAAAAATCTGTAGTAGTGCAGTGATCCAAAAGGAGGAAATTA	3360
Db	3301	TCACTTACTTTGCTTAAATAAAACACCGCAGTAGCAGTGATCCAAAAGGAGGAAACTG	3360
QY	3361	TCATCTGAAGTAAGACAGCAACGAGGGAATCAGAACACTTACCGGAGCCAGTGTGAT	3420
Db	3361	TCATCTGAATAAAGCAACGCGGAATCAGAACACTTACCGGAGCCAGTGTGAT	3420
QY	3421	TTTAAATCATAGTATGATTTTACACCCATATCCACTGTGTAAGAAAACATTAACAATTAGAG	3480
Db	3421	TTGAATCAGCGAAGACTTTCACA---AGCTCTCAGCACAGAAATCATTAACAGTAGAG	3477
QY	3481	ACTTCATTTATGGGGAATAAGCACATTTGAAGACACTGTGTAGTACACCAAGCATTTGAGAA	3540
Db	3478	CCTTCTCTCGTTGGGAATAAGCACCTTTGAAGACGCGGGCAGCACCTCAAAGTATTTGAGAG	3537
QY	3541	AATGACTTAAATTTCAACCAAGAAATTTTGTGTAAGAGAAATCAGAGAAATACAGCCCA	3600
Db	3538	AATGACTTCAAAATTTCCCAAGAGTTTGTGTAAGAGACTCAGAGAAACCAAGCCCA	3597
QY	3601	GAGAGGTTAGTATAGAAAGTTCAAACGAGCTCACATATGAAGATAGCTAGCCAAAGTTTC	3660
Db	3598	GAAAGTTGTTGTGGAAGGCTCAGCAAGCTCCACATCAAGATCCGACGCCAGAGCTTT	3657
QY	3661	AATAAGACACTACAAAGTGGCATTAAGTTCAATGAGTCAAGTCTTCAAGTCTTCAAGAGACA	3720
Db	3658	AACAGCGCACTACAGACGAGTGGCATCAGCTCCATGAGTCCGAGGCCCTCCGAGAGACA	3717
QY	3721	GTAGTGTAGTGTACAACTATGACACAGACTGTGGAAGCATGAGTACTGTGTAAGT	3780
Db	3718	GTGCTGTGCAACCCACAGCTATGACACGAGTGTGGAAGTCTGAGTACCCGGTGTGTAAGT	3777
QY	3781	ACTAAAATCTATTAAAGCAAGCCACTATTTTGAAGCCACAGCTCTAACCATCTCTCTCTCC	3840
Db	3778	ACTAAAATCTTTAAAGCAGGACCACTATTTGACACCAAGCTCTAACCATCTCTCTCTCC	3837
QY	3841	AAATCAAAATTCGGTGTCCCTGCTGCTCAGGTTCTTCTCATACGTTTCTTAGAAGACA	3900
Db	3838	AAATCGAACTCAGTGTCCCTGCTGCTCAGGTTCTTCTCACACCCCTTCTTAGAAGAGCC	3897
QY	3901	CAGTCCCTTAAAGCAGCCCTCTTATTTGCTACAAATTAAGAGTCTAGCAGATTTGTAACCT	3960
Db	3898	CAGTCCCTTAAAGCAGCCCTCTTATTTGCAACATTAAGAGTCTAGCAGATTTGTAACCT	3957
QY	3961	TACACAAGTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAAGA	4020

[illegible]

Db	5038	GTACAGTTTTTACAAATGCATGAAGAACGACAGAGCGCTGTGTGGCAATTCACCAATTTCAG	5099
Qy	5098	CCTATAGTTGATACATCTGCTGAATCTGA	5127
Db	5098	CCATAGTTGATGAGTCTGCTGAATCTGA	5127
RESULT 6			
LOCUS	CQ847976	3829 bp	DNA
DEFINITION	Sequence 281 from Patent WO2004063362.		linear
ACCESSION	CQ847976		
VERSION	CQ847976.1	GI:51469542	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1		
TITLE	Glover, D., Bell, G., Frenz, L. and Midgley, C.		
JOURNAL	Cell cycle progression proteins Patent: WO.2004063362-A 281 29-JUL-2004; ✓		
FEATURES	Cyclacel Limited (GB)		
source	Location/Qualifiers		
	1..3829		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="AB082530"		
ORIGIN			
	Query Match	74.7%;	Score 3829; DB 6; Length 3829;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 3829; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1299	GGCAACACAAATTTCTTCTCATTTCAATGACCCATCAATTACACTGCTTGGCCAAACCTTAAT	1358
Db	1	GGCAACACAAATTTCTTCTCATTTCAATGACCCATCAATTACACTGCTTGGCCAAACCTTAAT	60
Qy	1359	GAATATGGCTGCATCTTTTGATATCCCAAGGAAAGAGACTCGAGCCAGTCGACGCTT	1418
Db	61	GAATATGGCTGCATCTTTTGATATCCCAAGGAAAGAGACTCGAGCCAGTCGACGCTT	120
Qy	1419	GAATCTGTTAAAAACGGCTTCCATGAAATGAAGAAACGAGGACCTTAAGCCTTATAGTCTTCA	1478
Db	121	GAATCTGTTAAAAACGGCTTCCATGAAATGAAGAAACGAGGACCTTAAGCCTTATAGTCTTCA	180
Qy	1479	TTTAGACCAATTTATCAGAAAGCAATTCGACACACACACACACGGAATCAGTATCTCCG	1538
Db	181	TTTAGACCAATTTATCAGAAAGCAATTCGACACACACACACGGAATCAGTATCTCCG	240
Qy	1539	AGTTTCAGAAAGATATATTTATCCTTAAAGATACAGAGGAAGCTCTTTTAAATTAACCTTAG	1598
Db	241	AGTTTCAGAAAGATATATTTATCCTTAAAGATACAGAGGAAGCTCTTTTAAATTAACCTTAG	300
Qy	1599	AGATAGCCAAGTCCTTCAACATAAGAGAAATCTTGAATGGAATTTGGAATCTTTATAGGGAC	1658
Db	301	AGATAGCCAAGTCCTTCAACATAAGAGAAATCTTGAATGGAATTTGGAATCTTTATAGGGAC	360
Qy	1659	CATTCTTAAAGTGCCAAATGTAAATCTAAGAACTATAAAGATGAACAGTTACACAGGTT	1718
Db	361	CATTCTTAAAGTGCCAAATGTAAATCTAAGAACTATAAAGATGAACAGTTACACAGGTT	420
Qy	1719	TGTACGAGACTACTTTATTTTACAGGCCAGCAGTAAATATATATGCCAACCTGGATCT	1778
Db	421	TGTACGAGACTACTTTATTTTACAGGCCAGCAGTAAATATATATGCCAACCTGGATCT	480
Qy	1779	GGATTTTGGCAAGGCCAAACAGCTCAGCGTTGTAGGTTGCCAGTTTACAGAAATTTCTTCT	1838
Db	481	GGATTTTGGCAAGGCCAAACAGCTCAGCGTTGTAGGTTGCCAGTTTACAGAAATTTCTTCT	540
Qy	1839	TGAATCTGAAGAGGATGGCGAGGCTTACTAGAAAGATCTAGTAAAGGATATTGTTCCAGTG	1898

Db 541 TGAATCTGAAGAGGATGGCAAGGCTACTTAGAAGATCTAGTAAAGGATATTGTTTCAGTG 600
Qy 1899 GCTCAATGCTTCATCTGGAAATGAACCCGAAAGAGTCTTCAAAATAAATGGTTTATTGAC 1958
Db 601 GCTCAATGCTTCATCTGGAAATGAACCCGAAAGAGTCTTCAAAATAAATGGTTTATTGAC 660
Qy 1959 CACCCCTAGTCAACACTACTCTTTTATTTATTTGGAACACTTCTTCCGCCACCTCTCAGGAGT 2018
Db 661 CACCCCTAGTCAACACTACTCTTTTATTTATTTGGAACACTTCTTCCGCCACCTCTCAGGAGT 720
Qy 2019 TAAATGCTGGAAATGAGTGTATTTTCAGTGTCTCTTAAATCTTTGCTCTCTGAAATA 2078
Db 721 TAAATGCTGGAAATGAGTGTATTTTCAGTGTCTCTTAAATCTTTGCTCTCTGAAATA 780
Qy 2079 CCAAGATCACTTGTCTTAAATCTTACTGTGTTCTAGCTTGGACTATAGCAGAGATGGATTGGC 2138
Db 781 CCAAGATCACTTGTCTTAAATCTTACTGTGTTCTAGCTTGGACTATAGCAGAGATGGATTGGC 840
Qy 2139 TAGAGTCATCCTTCCAAAATTTTAACTGCAGCTACTGATGCTCGACACTCTATGCCAAC 2198
Db 841 TAGAGTCATCCTTCCAAAATTTTAACTGCAGCTACTGATGCTCGACACTCTATGCCAAC 900
Qy 2199 AAAACATTTAAGGCTATTATGAGAGCTAATGTGAAATCTTTTAAATTTGGGGAAATTGA 2258
Db 901 AAAACATTTAAGGCTATTATGAGAGCTAATGTGAAATCTTTTAAATTTGGGGAAATTGA 960
Qy 2259 GTTGTAGTGACCCAGCTACATGATAAAAAAAGAGATTTCTCTGAAAGCTCTTGATAT 2318
Db 961 GTTGTAGTGACCCAGCTACATGATAAAAAAAGAGATTTCTCTGAAAGCTCTTGATAT 1020
Qy 2319 CCTCGATGAGCAATGGAAGAGCCCAATCTTCATGCTCTCATTCAGATCAAAACCAGC 2378
Db 1021 CCTCGATGAGCAATGGAAGAGCCCAATCTTCATGCTCTCATTCAGATCAAAACCAGC 1080
Qy 2379 GTTATCCCACTTGGAGACAAGGGTTTGCTTCTCTGCTGAGATTTCTCTCCATCCAAA 2438
Db 1081 GTTATCCCACTTGGAGACAAGGGTTTGCTTCTCTGCTGAGATTTCTCTCCATCCAAA 1140
Qy 2439 AGAATTTCTATCTGAATGAAGAGGTTATGTAGCAAAACAAATGGAAAGTGGCAGAC 2498
Db 1141 AGAATTTCTATCTGAATGAAGAGGTTATGTAGCAAAACAAATGGAAAGTGGCAGAC 1200
Qy 2499 GGAATCAACTCCAAATATGTCATTTGATGAGGAACAACCTCAATGAAGCACTTACTAC 2558
Db 1201 GGAATCAACTCCAAATATGTCATTTGATGAGGAACAACCTCAATGAAGCACTTACTAC 1260
Qy 2559 TTACCGAAGCCTGTTGATGGTGATAACTATGTTTCGTTCGAGTAAACAAAGATTACAGCG 2618
Db 1261 TTACCGAAGCCTGTTGATGGTGATAACTATGTTTCGTTCGAGTAAACAAAGATTACAGCG 1320
Qy 2619 TCCTCAGCTACTCGCTATACCTTTATGACCACTAGTACACCATTAACACAGGCTG 2678
Db 1321 TCCTCAGCTACTCGCTATACCTTTATGACCACTAGTACACCATTAACACAGGCTG 1380
Qy 2679 CCATTTGTCGAAAGTACAGAAATATTATACAGAACTCTGTCGTAATGTTGTCACACGAG 2738
Db 1381 CCATTTGTCGAAAGTACAGAAATATTATACAGAACTCTGTCGTAATGTTGTCACACGAG 1440
Qy 2739 TTTGGATAAGTGGGAAGAAATTAATAAATGAAAGCAATCTTTTGGGCCCTTGGGAATAT 2798
Db 1441 TTTGGATAAGTGGGAAGAAATTAATAAATGAAAGCAATCTTTTGGGCCCTTGGGAATAT 1500
Qy 2799 CGGCTCATCAAAATGGGGTCTCAATTTGCTACAGGAAGAAACGTGATTTCCAGATATCT 2858
Db 1501 CGGCTCATCAAAATGGGGTCTCAATTTGCTACAGGAAGAAACGTGATTTCCAGATATCT 1560
Qy 2859 AAAATTTGCAAAACAGTGTGAAGTCTTCTCATCAGAGGACCTGTGTATATGTACTTGG 2918
Db 1561 AAAATTTGCAAAACAGTGTGAAGTCTTCTCATCAGAGGACCTGTGTATATGTACTTGG 1620
Qy 2919 GCTCATAGCTAAAAACCAAAACAGGCTGTGATATTTCTAAATGTCACAACTGGGATGCTGT 2978
Db 1621 GCTCATAGCTAAAAACCAAAACAGGCTGTGATATTTCTAAATGTCTCAACTCTGAAGCTTTGGCATC 1680

Qy 2979 GAGSCATAGTCGCAAAACATCTGTGGCCAGTGGTTCCAGATGATGTGGAACAACTCTGTAA 3038
Db 1681 GAGSCATAGTCGCAAAACATCTGTGGCCAGTGGTTCCAGATGATGTGGAACAACTCTGTAA 1740
Qy 3039 TGAACCTTTTCATCTATCCCAAGCACTCTAAAGTTTGAACCTCGAGGTCAACAGCTCTAGACA 3098
Db 1741 TGAACCTTTTCATCTATCCCAAGCACTCTAAAGTTTGAACCTCGAGGTCAACAGCTCTAGACA 1800
Qy 3099 TAATAGTGAAGTGAATCTGTGCCATCGAGTATGTCATATTCGAGATGACCGGTTTGG 3158
Db 1801 TAATAGTGAAGTGAATCTGTGCCATCGAGTATGTCATATTCGAGATGACCGGTTTGG 1860
Qy 3159 CAGCAGCTCTACTAGTATCAATTTTCTTGATATCAATGAAGATACAGAGCCAAATTTTA 3218
Db 1861 CAGCAGCTCTACTAGTATCAATTTTCTTGATATCAATGAAGATACAGAGCCAAATTTTA 1920
Qy 3219 TGACCGATCTGGACCCCAATAAAGGATAAAAAATTCATTCCTTTTCTTCTCTAGTAACT 3278
Db 1921 TGACCGATCTGGACCCCAATAAAGGATAAAAAATTCATTCCTTTTCTTCTCTAGTAACT 1980
Qy 3279 TGTGAAGATCGTATCTTAAATTCGCTTACTTTGSCCTTAAACAAAACATCTGTAGTAGCAG 3338
Db 1981 TGTGAAGATCGTATCTTAAATTCGCTTACTTTGSCCTTAAACAAAACATCTGTAGTAGCAG 2040
Qy 3339 TGATCCAAAAGCAGGGAATTTATCATCTGAAAAGTAAAGCAAGCAACAGGCGAAATCAGAAC 3398
Db 2041 TGATCCAAAAGCAGGGAATTTATCATCTGAAAAGTAAAGCAAGCAACAGGCGAAATCAGAAC 2100
Qy 3399 ACTTACGAGCCAGTGTGATTTTAAATCATAGTGTATTTTACAGCCATATCCACTGT 3458
Db 2101 ACTTACGAGCCAGTGTGATTTTAAATCATAGTGTATTTTACAGCCATATCCACTGT 2160
Qy 3459 ACAGAAACATTTAATTTAGAGACTTCATTTATGCGGAATTAAGCACATTTGAAGACACTGG 3518
Db 2161 ACAGAAACATTTAATTTAGAGACTTCATTTATGCGGAATTAAGCACATTTGAAGACACTGG 2220
Qy 3519 TAGTACACCAAGCAATGGAGAAAAATGACTTAAAAATTCACCAAGAAATTTTGGTACAGAAA 3578
Db 2221 TAGTACACCAAGCAATGGAGAAAAATGACTTAAAAATTCACCAAGAAATTTTGGTACAGAAA 2280
Qy 3579 TCACAGAGAAATACAGCCGAGAGAGGTTAGTAGTAAAGTTTCAAGGAGCTCACATAT 3638
Db 2281 TCACAGAGAAATACAGCCGAGAGAGGTTAGTAGTAAAGTTTCAAGGAGCTCACATAT 2340
Qy 3639 GAAGTACGTAGCCAAAGTTTCAATACAGACACTACACAAAGTGGCATAAGTTCAATCAG 3698
Db 2341 GAAGTACGTAGCCAAAGTTTCAATACAGACACTACACAAAGTGGCATAAGTTCAATCAG 2400
Qy 3699 CTCAAAGTCTTTCACGAGAGACAGTAGGTGTAGATGCTACAACTATGGAACAGACTGTGG 3758
Db 2401 CTCAAAGTCTTTCACGAGAGACAGTAGGTGTAGATGCTACAACTATGGAACAGACTGTGG 2460
Qy 3759 AAGATAGTACTGTGGTAAAGTACTTAAACTATTAAGCAAGCCACTATTTGACGCCACA 3818
Db 2461 AAGATAGTACTGTGGTAAAGTACTTAAACTATTAAGCAAGCCACTATTTGACGCCACA 2520
Qy 3819 GTCTAACCATCTCTCTCTCCAAATCAAATTCGGGTGCCCTGGGCTCCAGGTTCTTC 3878
Db 2521 GTCTAACCATCTCTCTCTCCAAATCAAATTCGGGTGCCCTGGGCTCCAGGTTCTTC 2580
Qy 3879 TCATACGCTTCTTAGAAGACACAGTCCCTTAAAGCACCTCTATTTGCTACAAATTAAG 3938
Db 2581 TCATACGCTTCTTAGAAGACACAGTCCCTTAAAGCACCTCTATTTGCTACAAATTAAG 2640
Qy 3939 TCTAGCAGATTTGTAATTTAGTTTACAGAGTTCTAGAGATGCTTTTGGCTATGCTACACT 3998
Db 2641 TCTAGCAGATTTGTAATTTAGTTTACAGAGTTCTAGAGATGCTTTTGGCTATGCTACACT 2700
Qy 3999 GAAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATC 4058
Db 2701 GAAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATC 2760

```
Qy 4059 TCCAGCAAAAGATGTGCTATTACTGATACCATCACCATGAAGGCCAACGATTTTGAGTC 4118
Db 2761 TCCAGCAAAAGATGTGCTATTACTGATACCATCACCATGAAGGCCAACGATTTTGAGTC 2820
Qy 4119 CAGATTAAACACCAAGCAGGTTTCATGAAGCCTTAAGTTATGATCATCATTTAGATAAAGAAGA 4178
Db 2821 CAGATTAAACACCAAGCAGGTTTCATGAAGCCTTAAGTTATGATCATCATTTAGATAAAGAAGA 2880
Qy 4179 TTTATTGAGTCCTATTAAATCAAAATACCTCCCAACGATTTTCCTCAGTGGCGGTCATGGT 4238
Db 2881 TTTATTGAGTCCTATTAAATCAAAATACCTCCCAACGATTTTCCTCAGTGGCGGTCATGGT 2940
Qy 4239 GTCCAGTGCCACATATGGGGGTTTCAGATGATTAATGTTGCTTCTCTCCCGTGGATAT 4298
Db 2941 GTCCAGTGCCACATATGGGGGTTTCAGATGATTAATGTTGCTTCTCTCCCGTGGATAT 3000
Qy 4299 AAATGATATATTTCCAGGTAAGAGGATATTCCTTATTTTCAGACAAAACATACACACACA 4358
Db 3001 AAATGATATATTTCCAGGTAAGAGGATATTCCTTATTTTCAGACAAAACATACACACACA 3060
Qy 4359 TGATGATCGAGGTGCAAGAGCATTTGCCCATGATGCGAGGAGTCTTCCATCTGGAACCTGG 4418
Db 3061 TGATGATCGAGGTGCAAGAGCATTTGCCCATGATGCGAGGAGTCTTCCATCTGGAACCTGG 3120
Qy 4419 AGGTCCTGTAATAAATCTTTTCACTTCTGCTAGCAGCAGATGATGCTTACGGAATAAT 4478
Db 3121 AGGTCCTGTAATAAATCTTTTCACTTCTGCTAGCAGCAGATGATGCTTACGGAATAAT 3180
Qy 4479 GAATTCATCATTCAGATGCTCTCTGTTTATTTAGAAAGTACAGAACACATGACATACA 4538
Db 3181 GAATTCATCATTCAGATGCTCTCTGTTTATTTAGAAAGTACAGAACACATGACATACA 3240
Qy 4539 GGAACATACAGATGAATGCTTATTTGTTGCTGTATTTGAAATCTTGGGTTTCCAGCC 4598
Db 3241 GGAACATACAGATGAATGCTTATTTGTTGCTGTATTTGAAATCTTGGGTTTCCAGCC 3300
Qy 4599 CAGCAACCACTGAGTGAATATGATGATTCATTCAGATTCAGATTCATATTCATATTCGA 4658
Db 3301 CAGCAACCACTGAGTGAATATGATGATTCATTCAGATTCAGATTCATATTCATATTCGA 3360
Qy 4659 TTGTTGAGCAGACTATCCATATCTTTAGAGTGGTTCCTTAAAGTTTTCGGGGAT 4718
Db 3361 TTGTTGAGCAGACTATCCATATCTTTAGAGTGGTTCCTTAAAGTTTTCGGGGAT 3420
Qy 4719 TTCTGATGATGATGGGGTGTCTCAAGAGGCTCAGCTAGCAGACCAAAAGCAGACA 4778
Db 3421 TTCTGATGATGATGGGGTGTCTCAAGAGGCTCAGCTAGCAGACCAAAAGCAGACA 3480
Qy 4779 ATTGTTACTAGGTTTAAACAAATTCAGATGATACCAATGTGCGGTATCTCTTCG 4838
Db 3481 ATTGTTACTAGGTTTAAACAAATTCAGATGATACCAATGTGCGGTATCTCTCTTCG 3540
Qy 4839 CAAGAAGTTCTAAGATTAGTATTAATTTGAGTGGTTCAGTTTCACTAAATGTCATGA 4898
Db 3541 CAAGAAGTTCTAAGATTAGTATTAATTTGAGTGGTTCAGTTTCACTAAATGTCATGA 3600
Qy 4899 GACTGGGCTTTTAAACAAATTAAGAGAGATGATCTTCAACAAATTTGATGACATATGCCCTTTA 4958
Db 3601 GACTGGGCTTTTAAACAAATTAAGAGAGATGATCTTCAACAAATTTGATGACATATGCCCTTTA 3660
Qy 4959 CTCTGAGGTTTCCCATTTGCTGTACATGACATTCAGACTTCCGTTGTCGAGGTTTCAT 5018
Db 3661 CTCTGAGGTTTCCCATTTGCTGTACATGACATTCAGACTTCCGTTGTCGAGGTTTCAT 3720
Qy 5019 ACAAGATTATTTCAAGATGATGATTTTCAAAATGATGAGAACGACAGGCTGTGTT 5078
Db 3721 ACAAGATTATTTCAAGATGATGATTTTCAAAATGATGAGAACGACAGGCTGTGTT 3780
Qy 5079 GGCACACACCAACCAAGCAACTATAGTTGATACATCTGCTGAATCCTCGA 5127
Db 3781 GGCACACACCAACCAAGCAACTATAGTTGATACATCTGCTGAATCCTCGA 3829
```

Query Match

74.7%; Score 3829; DB 9; Length 8213;

ORIGIN

```
RESULT 7
AB082530      8213 bp      mRNA      linear      PRI 04-JUL-2002
LOCUS      Homo sapiens mRNA for KIAA1999 protein.
DEFINITION      AB082530
ACCESSION      AB082530
VERSION      AB082530.1 GI:21693143
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Ohara,O., Nagase,T., Mitsui,G., Kohga,H., Kikuno,R., Hiraoka,S.,
Takahashi,Y., Kitajima,S., Saga,Y. and Koseki,H.
Characterization of size-fractionated cDNA libraries generated by
the in vitro recombination-assisted method
DNA Res. 9 (2), 47-57 (2002)
27051173
12056414
2 (bases 1 to 8213)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (26-MAR-2002) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
FEATURES
Location/Qualifiers
1..8213
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="bf00346"
/tissue-type="brain"
/notes="vector:pluescriptII SK plus"
1..8213
/genes="KIAA1999"
<1..3829
/genes="KIAA1999"
/notes="Start codon is not identified. bf00346 cDNA clone
for KIAA1999 has a 19-bp insertion after the position 1492
of the sequence of KIAA1999."
/codon_start=2
/evidence=not experimental
/product="KIAA1999 protein"
/protein_id="BAC02708.1"
/db_xref="GI:21693144"
/translation="ANTILPHSHSHLHCLPTLMNMAASFDIPKEKRLRASAALNCLK
RFHEMKRGPKYSLSLHDHIIQIAIATHQRDQYLRVQKDFILKDKTEALLINLRLD
QVQHKENLEWNNLIGTILKWPVNLRYNKDQLHRFVRLLYFYKPSKLYANLGL
DPAKAKQLTVGCGQTFELLESDEGQYLEDLVKDIVOMLNASSGMKPSRLONNGL
LTTLSOHYFLPIGTLSCHPHGVKLEKCSVFOCLNLCSLKNQDHLKLTUVSSLDYSR
DGLARVILSKILTAADACNLAYAKHLRVLLRANVEFFNNWGLLELLVTQLHDKNKTIS
SEALDILDEACEDKANHALIQMKPALSHGLDKGLLLRFLSPKGFSLNERYGYA
KQLEKWHREYNSKYDLIEEQLNEALTYRPKPDVGDNYVRNSQRLQRPVLYPIHLY
GOLVHHTGTCHLLEQVNIITELCNVRTPLDKWEEIKKLKASLWALNGTSSNWGLN
LLEQENVIPDLKLAQCEVLSIRGTCVVVLGIATKQKQCDILKCHNWDAVRHSRKH
LWFPVDDVEQLCNELSSIPSTLSASESTSSRHNSSESVSPSMKMFLEDDRFGSSST
LTFPDLINDEETFTFYDRSGPIKDKNSFPFASLKLKVNRLNLSLTPNKKHRSDDP
KGGKLSSEKTSNRRITLTPEPSVDFNHSDDFTPISTVQKTLQLETSFMGNKHIEDTG
STPSIGENDLKFTNFENHRENTRELRVSVESSTSHMKIRSQSFNTDITTSGIS
PGSSHTLPRAQSLKAPSIATIKSLADCNFSYTSRDAFGVATLKRLOQQRMHPSLSH
SEALAPAKDLVLTDTITMKANSFESRLTPSRFMKALSYSASLKEDLLSINQNTLOR
SSVRSMVSSATYGGSDDYIGLALPVDINDIFQVKDIPYQKNIIPHDDRGARAFAH
DAGGLGSGTGLKVNKSHLLRQQLMSLTIMNSISDASLEFSTEDTGLDHTDNCIL
YVCYIILGFTQPSNQLSAGTSDFDQIPYSDMCEQTIHNPLEVPSKFSGSCSG
VSGSGSASTKSTELLGKVTIPDDTMCRIKLRKVLRLVINLSSSVSTKCHETGGL
TIKEYPQTFEDDIDCLYSBVSHELLSHCTFRPCRRFIQELQVDFQFLQHEAEAVLAT
PKQKPIVDTSAES"
```

Best Local Similarity 100.0%; Pred. No. 0;				Matches 3829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	Db	QY	Db	QY	Db	QY	Db
1299	GGCAACCAATCTCTCTCATTCACATAGCCATCATTTACATGCTGCTGCCAACCTTAAT	1358	0;	0;	0;	0;	0;
1	GGCAACCAATCTCTCTCATTCACATAGCCATCATTTACATGCTGCTGCCAACCTTAAT	60					
1359	GAATATGGCTGCATCTCTTGTATATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCCCTT	1418					
61	GAATATGGCTGCATCTCTTGTATATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCCCTT	120					
1419	GAATGTTTAAACCGCTTCATGAATGAAGAAACGAGACCTTAAGCCCTTATAGTCTTCA	1478					
121	GAATGTTTAAACCGCTTCATGAATGAAGAAACGAGACCTTAAGCCCTTATAGTCTTCA	180					
1479	TTTAGACCACTATTCAGAAAGCAATTCGAACACACACAGGAAACCGGATCATGATCTCCG	1538					
181	TTTAGACCACTATTCAGAAAGCAATTCGAACACACACAGGAAACCGGATCATGATCTCCG	240					
1539	AGTTTCAGAAAGATATATTTATCTTAAAGATACAGAGGAAGCTCTTTTAAATTAACCTTAG	1598					
241	AGTTTCAGAAAGATATATTTATCTTAAAGATACAGAGGAAGCTCTTTTAAATTAACCTTAG	300					
1599	AGATAGCCAAAGCTCTTCAACATAAAGAGAACTTGAATGGAATTTGAAATCTTATAGGGAC	1658					
301	AGATAGCCAAAGCTCTTCAACATAAAGAGAACTTGAATGGAATTTGAAATCTTATAGGGAC	360					
1659	CATTCTTAAGTGGCCAAATGTAATCTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAG	1718					
361	CATTCTTAAGTGGCCAAATGTAATCTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAG	420					
1719	TGTACGAAGACTACTTTATTTTACAGGCCAGCAGTAAATTTATATGCCAACTTGATCT	1778					
421	TGTACGAAGACTACTTTATTTTACAGGCCAGCAGTAAATTTATATGCCAACTTGATCT	480					
1779	GAATTTGCAAGGCCAAACAGCTCACGTTGTAGGTTGCCAGTTTACAGAAATTTCTTCT	1838					
481	GAATTTGCAAGGCCAAACAGCTCACGTTGTAGGTTGCCAGTTTACAGAAATTTCTTCT	540					
1839	TGAATCTGAAGAGGATGGCCAGGCTACTTGAAGATCTAGTAAGATCTAGTAAGATCTAG	1898					
541	TGAATCTGAAGAGGATGGCCAGGCTACTTGAAGATCTAGTAAGATCTAGTAAGATCTAG	600					
1899	GCTCAATGCTTCATCTGGAATGAAACCCGAAAGAGTCTTCAAAATAATGTTTATTTGAC	1958					
601	GCTCAATGCTTCATCTGGAATGAAACCCGAAAGAGTCTTCAAAATAATGTTTATTTGAC	660					
1959	CACCTTAGTCAACACTACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	2018					
661	CACCTTAGTCAACACTACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	720					
2019	TAAATGCTGGAATGAGTATTTTACAGTGTCTCTTATCTTTGCTTCTTCTTCTTCTTCTT	2078					
721	TAAATGCTGGAATGAGTATTTTACAGTGTCTCTTATCTTTGCTTCTTCTTCTTCTTCTT	780					
2079	CCAAGATCACTTGTCTTAAACTTACTGTTTCTAGCTTGAAGTATAGCAGAGATGGATGGC	2138					
781	CCAAGATCACTTGTCTTAAACTTACTGTTTCTAGCTTGAAGTATAGCAGAGATGGATGGC	840					
2139	TAGAGTCACTCTTCCAAATTTTAACTGACGCTACTGATGCTGCTGACACTCTATGCAAC	2198					
841	TAGAGTCACTCTTCCAAATTTTAACTGACGCTACTGATGCTGCTGACACTCTATGCAAC	900					
2199	AAAACTTAAAGGATTTATTCAGAGCTAATCTTGAATCTTTTAAATTTGAGGATTTGA	2258					
901	AAAACTTAAAGGATTTATTCAGAGCTAATCTTGAATCTTTTAAATTTGAGGATTTGA	960					
2259	GTTGTTAGTGACCCAGCTACATGATATAAAACAAACGATTTCTCTGAGCTCTTGATAT	2318					
961	GTTGTTAGTGACCCAGCTACATGATATAAAACAAACGATTTCTCTGAGCTCTTGATAT	1020					
2319	CCTCGATGAAGCATGTGAAGACAAGGCCAATCTTTCATGCTCTCATTCAGATGAAACCCAGC	2378					
1021	CCTCGATGAAGCATGTGAAGACAAGGCCAATCTTTCATGCTCTCATTCAGATGAAACCCAGC	1080					
2379	GTTATCCCACTTTGAGACAAGGGTTTCTTCTCTGCTGAGATTTCTCTCATTTCCAAA	2438					
1081	GTTATCCCACTTTGAGACAAGGGTTTCTTCTCTGCTGAGATTTCTCTCATTTCCAAA	1140					
2439	AGGATTTTCTATCTGAATGAAGAGGTTATGTAGCAAAACAAATTTGGAAGTGGCACAG	2498					
1141	AGGATTTTCTATCTGAATGAAGAGGTTATGTAGCAAAACAAATTTGGAAGTGGCACAG	1200					
2499	GGATTAACAACCTCAAAATATTTGACTGTATGAGGAAACAACTCAATGAAGCACTTACTAC	2558					
1201	GGATTAACAACCTCAAAATATTTGACTGTATGAGGAAACAACTCAATGAAGCACTTACTAC	1260					
2559	TTACCGGAGGCTGTTGATGTTGATAAATCTGCTCGGAGTAAACCAAGATTTACAGCG	2618					
1261	TTACCGGAGGCTGTTGATGTTGATAAATCTGCTCGGAGTAAACCAAGATTTACAGCG	1320					
2619	TCCTCAAGCTTACCTGCTCTATACCTTTATGGAACAACTAGTACACCAATAAACAGGCTG	2678					
1321	TCCTCAAGCTTACCTGCTCTATACCTTTATGGAACAACTAGTACACCAATAAACAGGCTG	1380					
2679	CCATTTGTTGGAAGTACAGATATTTATTAAGAACTCTGCTGTAATGTTGCTACACAG	2738					
1381	CCATTTGTTGGAAGTACAGATATTTATTAAGAACTCTGCTGTAATGTTGCTACACAG	1440					
2739	TTTGATAGTGGGAGAAATTTAAACAACTGAAACATCTCTTGGGCTTGGGAAATAT	2798					
1441	TTTGATAGTGGGAGAAATTTAAACAACTGAAACATCTCTTGGGCTTGGGAAATAT	1500					
2799	CGGCTCATCAATTTGGGCTCTCAATTTGCTACAGGAAGAAACCGTATTTCCAGATATACT	2858					
1501	CGGCTCATCAATTTGGGCTCTCAATTTGCTACAGGAAGAAACCGTATTTCCAGATATACT	1560					
2859	AAAACTTGCAAAACAGTGTGAAGTTCTTTCCATCAGAGGACCTGTGTATATGTACTTGG	2918					
1561	AAAACTTGCAAAACAGTGTGAAGTTCTTTCCATCAGAGGACCTGTGTATATGTACTTGG	1620					
2919	GCTCATAGCTAAAAACCAACAGGCTGTGATATCTTAAATGCTCAACCTGGATGCTGT	2978					
1621	GCTCATAGCTAAAAACCAACAGGCTGTGATATCTTAAATGCTCAACCTGGATGCTGT	1680					
2979	GAGGCATAGTCGCAACATCTGTGGCCAGTGGTCCAGATGATGTGGAACAACTCTGTAA	3038					
1681	GAGGCATAGTCGCAACATCTGTGGCCAGTGGTCCAGATGATGTGGAACAACTCTGTAA	1740					
3039	TGAATCTTCACTATCCCAAGCACTCTAAGTTTGAATCGGAGTCAACAGCTCTAGACA	3098					
1741	TGAATCTTCACTATCCCAAGCACTCTAAGTTTGAATCGGAGTCAACAGCTCTAGACA	1800					
3099	TAAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTATATTTGGAGGATGACCGTTTGG	3158					
1801	TAAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTATATTTGGAGGATGACCGTTTGG	1860					
3159	CAGCAGCTCTACTAGTACATTTTCTTGTATCAATGAAGATACAGAGCCCAACATTTTAA	3218					
1861	CAGCAGCTCTACTAGTACATTTTCTTGTATCAATGAAGATACAGAGCCCAACATTTTAA	1920					
3219	TGACCGATCTGAGCCCAATGAAGATAAAAATTCATTCCTTCTTCTTCTTCTTCTTCTTCT	3278					
1921	TGACCGATCTGAGCCCAATGAAGATAAAAATTCATTCCTTCTTCTTCTTCTTCTTCTTCT	1980					
3279	TGTGAAGATCTGATCTTAAATTTGCTTACTTTGCTTAAACAAACATCTGTAGTAGCAG	3338					
1981	TGTGAAGATCTGATCTTAAATTTGCTTACTTTGCTTAAACAAACATCTGTAGTAGCAG	2040					
3339	TGATCAAAAGGAGGAAATTTATCATCTGAAAGTAAAGCAAGCAACAGCGCAATCAGAAC	3398					
2041	TGATCAAAAGGAGGAAATTTATCATCTGAAAGTAAAGCAAGCAACAGCGCAATCAGAAC	2100					
3399	ACTTACGAGGCCAGTGTGATTTTAAATCATAGTGTATTTTACACCATATCCACTGT	3458					
2101	ACTTACGAGGCCAGTGTGATTTTAAATCATAGTGTATTTTACACCATATCCACTGT	2160					

QY 3459 ACAGAAAACATTAACAATTAGAGACTTCAATTTATGCGGAATAGCAATTTGAGACACTGG 3518
DB 2161 ACAGAAAACATTAACAATTAGAGACTTCAATTTATGCGGAATAGCAATTTGAGACACTGG 2220
QY 3519 TAGTACACCAAGCATTTGGAGAAATGACTTAAATTTCCACCAAGATTTTGGTACAGAAA 3578
DB 2221 TAGTACACCAAGCATTTGGAGAAATGACTTAAATTTCCACCAAGATTTTGGTACAGAAA 2280
QY 3579 TCACAGAGAAAATACACAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACGAGCTCACATAT 3638
DB 2281 TCACAGAGAAAATACACAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACGAGCTCACATAT 2340
QY 3639 GAAGATACGTAGCCAAAGTTTCAATAACAGACACTACAAAGTGCCATAAGTTTCAATGAG 3698
DB 2341 GAAGATACGTAGCCAAAGTTTCAATAACAGACACTACAAAGTGCCATAAGTTTCAATGAG 2400
QY 3699 CTCAGTCCCTTCCACGAGAGACAGTAGGTAGTAGTCTCAACTATGGACACAGACTGG 3758
DB 2401 CTCAGTCCCTTCCACGAGAGACAGTAGGTAGTAGTCTCAACTATGGACACAGACTGG 2460
QY 3759 AAGCATGAGTACTGTGGTAAGTACTAAACTATTAAGACAGCCACTATTTGACGCCACA 3818
DB 2461 AAGCATGAGTACTGTGGTAAGTACTAAACTATTAAGACAGCCACTATTTGACGCCACA 2520
QY 3819 GTCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCCTCAGAGTTCTTC 3878
DB 2521 GTCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCCTCAGAGTTCTTC 2580
QY 3879 TCATACGCTTCTTAGAAGACAGCAGTCCCTTAAAGCAGCCCTCTATTGCTACATAAAG 3938
DB 2581 TCATACGCTTCTTAGAAGACAGCAGTCCCTTAAAGCAGCCCTCTATTGCTACATAAAG 2640
QY 3939 TCTAGCAGATGTAACTTTAGTTACACAAAGTTCTAGAGATGCTTTGGCTATGCTACACT 3998
DB 2641 TCTAGCAGATGTAACTTTAGTTACACAAAGTTCTAGAGATGCTTTGGCTATGCTACACT 2700
QY 3999 GAAAGACTACAGCAACCAAGAAATGCAATCCATCTCTCACTCTGAAGCTTTGGCATC 4058
DB 2701 GAAAGACTACAGCAACCAAGAAATGCAATCCATCTCTCACTCTGAAGCTTTGGCATC 2760
QY 4059 TCCAGCAAGATGTGCTATTACTGATACATACCATGAAAGGCCAACAGTTTGGATC 4118
DB 2761 TCCAGCAAGATGTGCTATTACTGATACATACCATGAAAGGCCAACAGTTTGGATC 2820
QY 4119 CAGATTAAACCAAGCAGGTTTCATGAAGCCTTAAGTTATGATCATCATTTAGATAAAGAGA 4178
DB 2821 CAGATTAAACCAAGCAGGTTTCATGAAGCCTTAAGTTATGATCATCATTTAGATAAAGAGA 2880
QY 4179 TTTATTGAGTCTTATTATCAAAATACCTCGAACGATCTTCTCAGTCCGCTCCATGGT 4238
DB 2881 TTTATTGAGTCTTATTATCAAAATACCTCGAACGATCTTCTCAGTCCGCTCCATGGT 2940
QY 4239 GTCCAGTCCCATATATGGGGTTCAGATGATTAATTTGCTTCTGCTCCCGGTGGATAT 4298
DB 2941 GTCCAGTCCCATATATGGGGTTCAGATGATTAATTTGCTTCTGCTCCCGGTGGATAT 3000
QY 4299 AAATGATATATTCAGAGTAAAGGATATTCCTTATTTTCAGACAAAACATACACACACA 4358
DB 3001 AAATGATATATTCAGAGTAAAGGATATTCCTTATTTTCAGACAAAACATACACACACA 3060
QY 4359 TGATGATCGAGGTGCAAGAGCATTTGGCCATGATGCGAGAGGTCTTCCATCTGGAACCTGG 4418
DB 3061 TGATGATCGAGGTGCAAGAGCATTTGGCCATGATGCGAGAGGTCTTCCATCTGGAACCTGG 3120
QY 4419 AGGTCTTGTAATAAATCTTTTCTAGTCTGCTACGACAGAGATGATCTTACGGAATAAT 4478
DB 3121 AGGTCTTGTAATAAATCTTTTCTAGTCTGCTACGACAGAGATGATCTTACGGAATAAT 3180
QY 4479 GAAATTCATTCATTCAGATGCTCTCTGTTTATAGAAAGTACAGACACTGACATACA 4538
DB 3181 GAAATTCATTCATTCAGATGCTCTCTGTTTATAGAAAGTACAGACACTGACATACA 3240

QY 4539 GGAACATACAGATGATRACTGCCCTTTATTTGTTCTGTATTGAAATTTCTGGGTTCCAGCC 4598
DB 3241 GGAACATACAGATGATRACTGCCCTTTATTTGTTCTGTATTGAAATTTCTGGGTTCCAGCC 3300
QY 4599 CAGCAACCAACTGAGTGCATATATGATGATCATTCAGACTTTCAAGATATTCATATTTCTGA 4658
DB 3301 CAGCAACCAACTGAGTGCATATATGATGATCATTCAGACTTTCAAGATATTCATATTTCTGA 3360
QY 4659 TTGGTGTGAGCAGACTATCCATAATCCCTTTAGAAAGTGGTTCCTCTAAGTTTTCGGGAT 4718
DB 3361 TTGGTGTGAGCAGACTATCCATAATCCCTTTAGAAAGTGGTTCCTCTAAGTTTTCGGGAT 3420
QY 4719 TTCTGGATGCAATGATGGGGTGTCTCAAGAAAGGCTCAGCTAGCAGCAACCAAGACACAGA 4778
DB 3421 TTCTGGATGCAATGATGGGGTGTCTCAAGAAAGGCTCAGCTAGCAGCAACCAAGACACAGA 3480
QY 4779 ATTCTTACTAGTGTCTTAAACCAATTCAGATGATACACCAATGTCGGTATCTCTTCG 4838
DB 3481 ATTCTTACTAGTGTCTTAAACCAATTCAGATGATACACCAATGTCGGTATCTCTTCG 3540
QY 4839 CAAAGAAAGTTCTTAAGATTAGTCAATTAATTTGAGTAGTTTCAACTTAAATGTCATGA 4898
DB 3541 CAAAGAAAGTTCTTAAGATTAGTCAATTAATTTGAGTAGTTTCAACTTAAATGTCATGA 3600
QY 4899 GACTGGGCTTTTAAACAATTAAGAGAAAGTATCTCTCAAACTTTGATGACATATGCTTTA 4958
DB 3601 GACTGGGCTTTTAAACAATTAAGAGAAAGTATCTCTCAAACTTTGATGACATATGCTTTA 3660
QY 4959 CTCTGAGGTTTCCCATTTGCTGTCACATGACATTCAGACTTCCGTTCCGAGGTTTCAAT 5018
DB 3661 CTCTGAGGTTTCCCATTTGCTGTCACATGACATTCAGACTTCCGTTCCGAGGTTTCAAT 3720
QY 5019 ACAAGAAATTTTCAAGATGTACAGATTTCTTACAATGATGATGATGATGATGATGATGAT 5078
DB 3721 ACAAGAAATTTTCAAGATGTACAGATTTCTTACAATGATGATGATGATGATGATGAT 3780
QY 5079 GGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTGAATCCTGA 5127
DB 3781 GGCAACACCAACCAAGCAACCTATAGTTGATGATGATGATGATGATGATGATGAT 3829

RESULT 8

BD185164
LOCUS BD185164 8232 bp DNA linear PAT 17-JUN-2003
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD185164
VERSION BD185164.1 GI:31877364
KEYWORDS JP 2002345493-A/7.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8232)
AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
TITLE Novel genes and proteins encoded by the genes
JOURNAL Patent: JP 2002345493-A 7 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002345493-A/7
PD 03-DEC-2002
PF 26-FEB-2002 JP 2002049046
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09, C07K14/47, C07K14/54, C12N15/00
CC Novel genes and proteins encoded by the genes FH Key

FEATURES : FT CDS (1473).. (3845).

Location/Qualifiers
1..8232
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 74.1%; Score 3800; DB 6; Length 8232; Best Local Similarity 99.5%; Pred. No. 0; Matches 3829; Conservative 0; Mismatches 0; Indels 19; Gaps 1;									
QY	1299	GGCAAAACAATTCCTTCATTCACATAGCCATCATTTACATCTGCTGCCAACCCCTAAT	1358						
DB	1	GGCAAAACAATTCCTTCATTCACATAGCCATCATTTACATCTGCTGCCAACCCCTAAT	60						
QY	1359	GAATATGGCTGCATCCTTTGATATCCCAAGGAAAGAGACTGCGAGCGAGTGCAGCCTT	1418						
DB	61	GAATATGGCTGCATCCTTTGATATCCCAAGGAAAGAGACTGCGAGCGAGTGCAGCCTT	120						
QY	1419	GAATCTTTAAACGCTTCCATGAATGAAGAACGAGACCTTAAGCCTTATAGTCTTCA	1478						
DB	121	GAATCTTTAAACGCTTCCATGAATGAAGAACGAGACCTTAAGCCTTATAGTCTTCA	180						
QY	1479	TTTAGACCAATTAATTCAGAAAGCAATTTGCAACACACCAAGAAACGGGATCAGTATCTCG	1538						
DB	181	TTTAGACCAATTAATTCAGAAAGCAATTTGCAACACACCAAGAAACGGGATCAGTATCTCG	240						
QY	1539	AGTTGAGAAAGATATATTTATCTTAAGATACAGAGGAGCTCTTTTAAATTAACCTTAG	1598						
DB	241	AGTTGAGAAAGATATATTTATCTTAAGATACAGAGGAGCTCTTTTAAATTAACCTTAG	300						
QY	1599	AGATAGCCAGTCTTCAACATAAAGAGAACTTTGAATGGAATTTGAAATCTTATAGGGAC	1658						
DB	301	AGATAGCCAGTCTTCAACATAAAGAGAACTTTGAATGGAATTTGAAATCTTATAGGGAC	360						
QY	1659	CATTCCTTAAGTGGCCAAATGTAATCTTAAGAACTATAAGATGAACAGTTACACAGTT	1718						
DB	361	CATTCCTTAAGTGGCCAAATGTAATCTTAAGAACTATAAGATGAACAGTTACACAGTT	420						
QY	1719	TGTACGAAGACTATTTATTTTACAGCCAGCAGTAATATATATAGCCAACTTGATCT	1778						
DB	421	TGTACGAAGACTATTTATTTTACAGCCAGCAGTAATATATATAGCCAACTTGATCT	480						
QY	1779	GGATTTTGCCAGCCAAACAGCTCAGGTTGTAGTTGCGAGTTTACAGAAATTTCTTCT	1838						
DB	481	GGATTTTGCCAGCCAAACAGCTCAGGTTGTAGTTGCGAGTTTACAGAAATTTCTTCT	540						
QY	1839	TGAATCTGAAGAGGATGGGCAAGGCTACTTAGAAGATCTAGTAAAGGATATTTTCAGTG	1898						
DB	541	TGAATCTGAAGAGGATGGGCAAGGCTACTTAGAAGATCTAGTAAAGGATATTTTCAGTG	600						
QY	1899	GCTCAATGCTTCTGTGGAATGAACCCGAAAGAGTCTTCAAAATAATGGTTTATGAC	1958						
DB	601	GCTCAATGCTTCTGTGGAATGAACCCGAAAGAGTCTTCAAAATAATGGTTTATGAC	660						
QY	1959	CACCTTAGTCAACACTACTTTTATTTATTTGGAACACTTTCTTGCCACCCTCATGGAGT	2018						
DB	661	CACCTTAGTCAACACTACTTTTATTTATTTGGAACACTTTCTTGCCACCCTCATGGAGT	720						
QY	2019	TAAATGCTGGAAAAATGCAAGTGTATTTCAAGTGTCTCTTAAATCTTTGCTCTTGAATA	2078						
DB	721	TAAATGCTGGAAAAATGCAAGTGTATTTCAAGTGTCTCTTAAATCTTTGCTCTTGAATA	780						
QY	2079	CCAAGATCACTTGCTAAACCTTACTGTTCTAGCTTGGACTATAGCAGAGATGGATGGC	2138						
DB	781	CCAAGATCACTTGCTAAACCTTACTGTTCTAGCTTGGACTATAGCAGAGATGGATGGC	840						
QY	2139	TAGAGTCATCTTTTCCAAAAATTTAACTGCAGCTACTGATGCTGCAGACTCTATGCAAC	2198						
DB	841	TAGAGTCATCTTTTCCAAAAATTTAACTGCAGCTACTGATGCTGCAGACTCTATGCAAC	900						
QY	2199	AAAAATTTAAGGGTATTTATGAGAGCTAATGTTGAATCTTTTAAATTTGGGAAATGA	2258						
DB	901	AAAAATTTAAGGGTATTTATGAGAGCTAATGTTGAATCTTTTAAATTTGGGAAATGA	960						
QY	2259	GTTGTTAGTACCAGCTACATGATAAAACAAACGATTTCTCTGAAGCTCTTGATAT	2318						
DB	961	GTTGTTAGTACCAGCTACATGATAAAACAAACGATTTCTCTGAAGCTCTTGATAT	1020						
QY	2319	CCTCGATGAAGCATGTGAAGACAGGCGCAATCTTCTGCTCTCATTCATTCAGATGAACCGAC	2378						
DB	1021	CCTCGATGAAGCATGTGAAGACAGGCGCAATCTTCTGCTCTCATTCATTCAGATGAACCGAC	1080						
QY	2379	GTTATCCCACTTGTGAGACAAGGGTTTCTTCTGCTGAGATTTCTCTCAATTCCTCAAA	2438						
DB	1081	GTTATCCCACTTGTGAGACAAGGGTTTCTTCTGCTGAGATTTCTCTCAATTCCTCAAA	1140						
QY	2439	AGGATTTTCTTATCTGAATGAAGAGGTTATGTAGCAAAACAATTTGGAAGAGTGCACAG	2498						
DB	1141	AGGATTTTCTTATCTGAATGAAGAGGTTATGTAGCAAAACAATTTGGAAGAGTGCACAG	1200						
QY	2499	GGAAATACAACCTCAAAATATGTTGATTTGATTGAGGAACAACCTCAATGAAGCACTTACTAC	2558						
DB	1201	GGAAATACAACCTCAAAATATGTTGATTTGATTGAGGAACAACCTCAATGAAGCACTTACTAC	1260						
QY	2559	TTACCGGAAGCCTGTTGATGATGATACTATGTTTCTGTCGAGTAACCAAGATTTACAGG	2618						
DB	1261	TTACCGGAAGCCTGTTGATGATGATACTATGTTTCTGTCGAGTAACCAAGATTTACAGG	1320						
QY	2619	TCCTCAGCTCTACCTGCTTATACACCTTTATGGACAACCTAGTACACCAATGAAGAGGCTG	2678						
DB	1321	TCCTCAGCTCTACCTGCTTATACACCTTTATGGACAACCTAGTACACCAATGAAGAGGCTG	1380						
QY	2679	CCATTTGTTGGAAGTACAGAAATATTTATTAAGAACTCTGTCGTAATGTTCTGTAACCCAG	2738						
DB	1381	CCATTTGTTGGAAGTACAGAAATATTTATTAAGAACTCTGTCGTAATGTTCTGTAACCCAG	1440						
QY	2739	TTTGGATATGTTGGGAAGAAATTTAAAAAAGTGAAGCACTCTTTGGGCTTG-----	2790						
DB	1441	TTTGGATATGTTGGGAAGAAATTTAAAAAAGTGAAGCACTCTTTGGGCTTGTTGTA	1500						
QY	2791	-----GGAATATCGGCTCATCAAAATTTGGGCTCTCAATTTGCTACAGGAAGAA	2839						
DB	1501	TAAACACTTCTGGAATATCGGCTCATCAAAATTTGGGCTCTCAATTTGCTACAGGAAGAA	1560						
QY	2840	ACGTGATTCAGATATACATAAACTTTGCAAAACAGTGTGAAGTTCTTTCCATCAGAGGA	2899						
DB	1561	ACGTGATTCAGATATACATAAACTTTGCAAAACAGTGTGAAGTTCTTTCCATCAGAGGA	1620						
QY	2900	CCTGTGTATATGATCTTGGGCTCATAGCTTAAACCAACAGGCTGTGATATCTTAAAT	2959						
DB	1621	CCTGTGTATATGATCTTGGGCTCATAGCTTAAACCAACAGGCTGTGATATCTTAAAT	1680						
QY	2960	GTCACAACTGGGATGCTGTGAGGCAATGTCGCAAAACATCTGTGGCAGTGGTTCCAGATG	3019						
DB	1681	GTCACAACTGGGATGCTGTGAGGCAATGTCGCAAAACATCTGTGGCAGTGGTTCCAGATG	1740						
QY	3020	ATGTGGAACAACCTCTGTAAATGAACCTTATATCCCAAGCACTCTTAAGTTGAACCTCG	3079						
DB	1741	ATGTGGAACAACCTCTGTAAATGAACCTTATATCCCAAGCACTCTTAAGTTGAACCTCG	1800						
QY	3080	AGTCAACAGCTCTAGACATATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATAT	3139						
DB	1801	AGTCAACAGCTCTAGACATATAGTGAAGTGAATCTGTGCCATCGAGTATGTTTCATAT	1860						
QY	3140	TGAGAGATCAACCGGTTTGGCAGCAGCTCTACTAGTACATTTTCTTGTATCAATGAAG	3199						
DB	1861	TGAGAGATCAACCGGTTTGGCAGCAGCTCTACTAGTACATTTTCTTGTATCAATGAAG	1920						
QY	3200	ATACAGAGCCAACTTTTATGACCGGATCTGGACCCCAATAAGGATATAAATTTCAATCCCT	3259						
DB	1921	ATACAGAGCCAACTTTTATGACCGGATCTGGACCCCAATAAGGATATAAATTTCAATCCCT	1980						
QY	3260	TCTTTGCTTCTAGTAAACCTTGTGAAGATCGTATCTTAAATTCGCTTACTTGGCTTAA	3319						
DB	1981	TCTTTGCTTCTAGTAAACCTTGTGAAGATCGTATCTTAAATTCGCTTACTTGGCTTAA	2040						
QY	3320	AAAAACATCTGTAGTACAGTATCCAAAAGGAGGGAATTTATCATCTGAAAGTAAACAA	3379						
DB	2041	AAAAACATCTGTAGTACAGTATCCAAAAGGAGGGAATTTATCATCTGAAAGTAAACAA	2100						
QY	3380	GCAACAGGCGAATCAGAACTTTACGAGCCAGTGTGATTTTAAATCATCATGATGATTT	3439						

Db 2101 GCACAGCGGGAATCAGAACACTTACGAGCCAGCTGTTGATTTTAAATCATAGTGTGATG 2160
Qy 3440 TTACACCCATATCCACTGTGTACAGAAAAATTAACAATTAGAGACTTTCATTTATGGGGAATA 3499
Db 2161 TTACACCCATATCCACTGTGTACAGAAAAATTAACAATTAGAGACTTTCATTTATGGGGAATA 2220
Qy 3500 AGCACAATTGAAGACACTGGTGTAGTACACAGCAATTTGGAGAAAAATGACTTAAATTTACCA 3559
Db 2221 AGCACAATTGAAGACACTGGTGTAGTACACAGCAATTTGGAGAAAAATGACTTAAATTTACCA 2280
Qy 3560 AGAATTTTGGTACAGAGAAATCAGAGAAAAATACAAAGCCGAGAGAGGTTAGTAGTAGAAA 3619
Db 2281 AGAATTTTGGTACAGAGAAATCAGAGAAAAATACAAAGCCGAGAGAGGTTAGTAGTAGAAA 2340
Qy 3620 GTTCAACGAGCTCACATATGAAGATACGCTAGCCTAAAGTTTCAATACAGACACTACAA 3679
Db 2341 GTTCAACGAGCTCACATATGAAGATACGCTAGCCTAAAGTTTCAATACAGACACTACAA 2400
Qy 3680 GTGGCATAAGTTCAATGAGCTCAAGTCTTCCAGAGACAGTAGGTTGATGCTACAA 3739
Db 2401 GTGGCATAAGTTCAATGAGCTCAAGTCTTCCAGAGACAGTAGGTTGATGCTACAA 2460
Qy 3740 CTATGGACACAGACTGTGGAGAGCATGACTGTGTGTTAGTAACTATAAATTTAAAGCAA 3799
Db 2461 CTATGGACACAGACTGTGGAGAGCATGACTGTGTGTTAGTAACTATAAATTTAAAGCAA 2520
Qy 3800 GCCACTATTTGACGCGCACAGCTTAACCATCTGTCTCTCCAAATCAAAATCGGTGTC 3859
Db 2521 GCCACTATTTGACGCGCACAGCTTAACCATCTGTCTCTCCAAATCAAAATCGGTGTC 2580
Qy 3860 TGGTGCTTCCAGGTTCTTCTCATAGCTTCTAGAGACAGCTCCCTTAAAGCACCCCT 3919
Db 2581 TGGTGCTTCCAGGTTCTTCTCATAGCTTCTAGAGACAGCTCCCTTAAAGCACCCCT 2640
Qy 3920 CTATTTGTACAAATAAAGTCTAGCAGATTTCTAATCTTGTAGTACAGAGTTCTAGAGATG 3979
Db 2641 CTATTTGTACAAATAAAGTCTAGCAGATTTCTAATCTTGTAGTACAGAGTTCTAGAGATG 2700
Qy 3980 CTTTGTGGCTATGTCTACACTGAAAAGACTACAGCAACAAAGATGCATCCATCTTATCTC 4039
Db 2701 CTTTGTGGCTATGTCTACACTGAAAAGACTACAGCAACAAAGATGCATCCATCTTATCTC 2760
Qy 4040 ACTCTGAAGCTTTGGGCTATCTCAGCAAAAGATGTGCTATTTAATCTGATACCAATCAACATGA 4099
Db 2761 ACTCTGAAGCTTTGGGCTATCTCAGCAAAAGATGTGCTATTTAATCTGATACCAATCAACATGA 2820
Qy 4100 AGGCCAACAGTTTGTAGTCCAGATTAACACCAAGCAGGTTTCATGAAGCCCTTAAGTTATG 4159
Db 2821 AGGCCAACAGTTTGTAGTCCAGATTAACACCAAGCAGGTTTCATGAAGCCCTTAAGTTATG 2880
Qy 4160 CATCATTAGATAAAGAGATTTATGTAGTCTTATTAATCAAAATACCTGCAACGATCTT 4219
Db 2881 CATCATTAGATAAAGAGATTTATGTAGTCTTATTAATCAAAATACCTGCAACGATCTT 2940
Qy 4220 CCTCAGTCCGCTCCATGGTGTCCAGTGCCACATATGGGGTTCAGATGATTACATTTGGTC 4279
Db 2941 CCTCAGTCCGCTCCATGGTGTCCAGTGCCACATATGGGGTTCAGATGATTACATTTGGTC 3000
Qy 4280 TTGCTCTCCCGTGGATATAAATCATATATTCAGGTAAAGGATTTCCCTATTTTCAGA 4339
Db 3001 TTGCTCTCCCGTGGATATAAATCATATATTCAGGTAAAGGATTTCCCTATTTTCAGA 3060
Qy 4340 CAAAAACATACACACATCATGATCATGAGTGCAAGAGCAATTTGCCCATGATGATGAGAG 4399
Db 3061 CAAAAACATACACACATCATGATCATGAGTGCAAGAGCAATTTGCCCATGATGATGAGAG 3120
Qy 4400 GTCTTCCATCTGGAATCTGGAGGTTCTGTAAATAATTTCTTTCACTTGTCTAGCAGCAGAGA 4459
Db 3121 GTCTTCCATCTGGAATCTGGAGGTTCTGTAAATAATTTCTTTCACTTGTCTAGCAGCAGAGA 3180
Qy 4460 TGAGTCTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTGTTTGAAGAAGTA 4519
Db 3181 TGAGTCTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTGTTTGAAGAAGTA 3240

Qy 4520 CAGAAGACACTGGACTACAGGAACATACAGATGATAACTGCCTTTATTTGTTGTTGTTATG 4579
Db 3241 CAGAAGACACTGGACTACAGGAACATACAGATGATAACTGCCTTTATTTGTTGTTGTTATG 3300
Qy 4580 AAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTC 4639
Db 3301 AAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTC 3360
Qy 4640 AAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCATAATCTTTAGAGTGGTTC 4699
Db 3361 AAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCATAATCTTTAGAGTGGTTC 3420
Qy 4700 CCTCTAAGTTTTCGGGGATTTCTGATGTCAGTGTGGGTGTCTCAAGAAAGCTCAGCTA 4759
Db 3421 CCTCTAAGTTTTCGGGGATTTCTGATGTCAGTGTGGGTGTCTCAAGAAAGCTCAGCTA 3480
Qy 4760 GCAGACCAAAAGACAGAAATTTGTACTAGTGTTTAAACAATTCAGATGATACACCAA 4819
Db 3481 GCAGACCAAAAGACAGAAATTTGTACTAGTGTTTAAACAATTCAGATGATACACCAA 3540
Qy 4820 TGTCCGCTTACTCTTTCGCAAGAGTTCTAAGTATAGTCAATTAATTTAGTAGTTTCTAG 4879
Db 3541 TGTCCGCTTACTCTTTCGCAAGAGTTCTAAGTATAGTCAATTAATTTAGTAGTTTCTAG 3600
Qy 4880 TTTCAACTAAATGTCTGAGACTGGGCTTTTAAACAATTAAGGAGAGTATCTCTCAACAT 4939
Db 3601 TTTCAACTAAATGTCTGAGACTGGGCTTTTAAACAATTAAGGAGAGTATCTCTCAACAT 3660
Qy 4940 TTGATGACATATGCTTACTCTGAGGTTTCCCAATTTGCTGTCACACTGCACATTCAGAC 4999
Db 3661 TTGATGACATATGCTTACTCTGAGGTTTCCCAATTTGCTGTCACACTGCACATTCAGAC 3720
Qy 5000 TTCCGCTCTGGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTCAAAATGCATG 5059
Db 3721 TTCCGCTCTGGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTCAAAATGCATG 3780
Qy 5060 AAGAAGCAGAGGCTGTGTGGCAACACACCAAGCAACCTATAGTTGATACATCTGCTG 5119
Db 3781 AAGAAGCAGAGGCTGTGTGGCAACACACCAAGCAACCTATAGTTGATACATCTGCTG 3840
Qy 5120 AATCCTGA 5127
Db 3841 AATCCTGA 3848

RESULT 9

BC051729

LOCUS

DEFINITION

IMAGE:5787163), partial cds.

ACCESSION

BC051729

VERSION

BC051729.1 GI:30704351

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

5440 bp mRNA linear PRI 27-JAN-2004
Homo sapiens hypothetical protein MGC39830, mRNA (cdna clone)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5440)
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Db 1021 AAACATCTGTGCCAGTGGTCCAGATGATGTGGAAACAACTCTGTAATGAACCTTTCACTCT 1080
Qy 3052 ATCCCAAGCACTCTAAAGTTTGAACCTCGAGGTCAACCAAGCTCTAGACATAAATAGTGAAGT 3111
Db 1081 ATCCCAAGCACTCTAAAGTTTGAACCTCGAGGTCAACCAAGCTCTAGACATAAATAGTGAAGT 1140
Qy 3112 GAATCTGTGCCATCGAGTATGTTCTATATGAGAGATGACCGGTTGGCAGCAGCTCTACT 3171
Db 1141 GAATCTGTGCCATCGAGTATGTTCTATATGAGAGATGACCGGTTGGCAGCAGCTCTACT 1200
Qy 3172 AGTACATTTTCCCTTGATATCAATCAAGATACAGAGCCAAATTTTATGACCGATCTGGA 3231
Db 1201 AGTACATTTTCCCTTGATATCAATCAAGATACAGAGCCAAATTTTATGACCGATCTGGA 1260
Qy 3232 CCCATAAAGGATAAAAATTCATTCCTCTTCTTCTTCTAGTAAACTTGTGAAGATCGT 3291
Db 1261 CCCATAAAGGATAAAAATTCATTCCTCTTCTTCTTCTAGTAAACTTGTGAAGATCGT 1320
Qy 3292 ATCTTAAATTCGCTTACCTTTCGCTTAAACAAAAAATCGTAGTACAGTGATCCAAAAGGA 3351
Db 1321 ATCTTAAATTCGCTTACCTTTCGCTTAAACAAAAAATCGTAGTACAGTGATCCAAAAGGA 1380
Qy 3352 GGGAAATTTATCATCTGAAAGTAAGACAGCAAGCAAGCGCAATCAGAACACTTTACGGAGCCC 3411
Db 1381 GGGAAATTTATCATCTGAAAGTAAGACAGCAAGCAAGCGCAATCAGAACACTTTACGGAGCCC 1440
Qy 3412 AGTGTGATTTTAAATCATAGTAGTAAATTTTACACCCATATCCACTGTACAGAAACATTA 3471
Db 1441 AGTGTGATTTTAAATCATAGTAGTAAATTTTACACCCATATCCACTGTACAGAAACATTA 1500
Qy 3472 CAATTAGAGACTTCATTTATGGGAATAAGACACTTGAAGACACTGGTAGTACACCAAGC 3531
Db 1501 CAATTAGAGACTTCATTTATGGGAATAAGACACTTGAAGACACTGGTAGTACACCAAGC 1560
Qy 3532 ATTGGAGAAATGACTTAAATTCACCAAGAAATTTTGGTACAGAGAAATCAGAGAGAAAT 3591
Db 1561 ATTGGAGAAATGACTTAAATTCACCAAGAAATTTTGGTACAGAGAAATCAGAGAGAAAT 1620
Qy 3592 ACAAGCCGAGAGAGGTAGTAGTAGAAGTTTCAACGAGCTCACATATAGAGATACGTAGC 3651
Db 1621 ACAAGCCGAGAGAGGTAGTAGTAGAAGTTTCAACGAGCTCACATATAGAGATACGTAGC 1680
Qy 3652 CAAAGTTTCAATACAGACACTACAAAGTGGCATAAGTTCAATGAGCTCAAGTCCCTTCA 3711
Db 1681 CAAAGTTTCAATACAGACACTACAAAGTGGCATAAGTTCAATGAGCTCAAGTCCCTTCA 1740
Qy 3712 CGAGAGACGTAGGTGTAGATGCTACAACTATGAGACACAGACTGTGGAAGCATGAGTACT 3771
Db 1741 CGAGAGACGTAGGTGTAGATGCTACAACTATGAGACACAGACTGTGGAAGCATGAGTACT 1800
Qy 3772 GTGTAAAGTACTAAACCTATTAAAGCAAGCCACTATTGAGCCACACAGCTTAAACCATCTG 3831
Db 1801 GTGTAAAGTACTAAACCTATTAAAGCAAGCCACTATTGAGCCACACAGCTTAAACCATCTG 1860
Qy 3832 TCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGGCTCCAGGTTCTCTCATACGCTTCCCT 3891
Db 1861 TCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGGCTCCAGGTTCTCTCATACGCTTCCCT 1920
Qy 3892 AGAAGACACAGTCCCTTAAAGCAACCTCTATTGCTCAAAATTAAGCTTAGCAGATGTG 3951
Db 1921 AGAAGACACAGTCCCTTAAAGCAACCTCTATTGCTCAAAATTAAGCTTAGCAGATGTG 1980
Qy 3952 AACCTTTAGTTACAAAGTCTTAGAGATGCTTTGGCTATGCTACACTGAAAGACTACAG 4011
Db 1981 AACCTTTAGTTACAAAGTCTTAGAGATGCTTTGGCTATGCTACACTGAAAGACTACAG 2040
Qy 4012 CAACAAAGATGCAATCCATCTTATCTCACTGAAAGCTTTGGCATCTCCAGCAAAAGAT 4071
Db 2041 CAACAAAGATGCAATCCATCTTATCTCACTGAAAGCTTTGGCATCTCCAGCAAAAGAT 2100
Qy 4072 GTGCTATTTTACTGATACCATCAACATGAAGCCCAAGTTTGTGATTCAGATTAACACCA 4131

RESULT 10
AK173326

Db 2101 GTGCTATTTTACTGATACCATCCATGAAGCCCAACAGTTTGTAGTCCAGATTAACACCA 2160
Qy 4132 AGCAGGTTTCATGAAGCCCTTAAGTTTATGATCATTTAGATAAAGAGATTTTATTTAGTCTCT 4191
Db 2161 AGCAGGTTTCATGAAGCCCTTAAGTTTATGATCATTTAGATAAAGAGATTTTATTTAGTCTCT 2220
Qy 4192 ATTAATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCTCAGTGGTCCATGAGTGCACCA 4251
Db 2221 ATTAATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCTCAGTGGTCCATGAGTGCACCA 2280
Qy 4252 TATGGGGTTTCAGATGATTAATTTGCTTCTCCGTTGGATATAAATGATATATTC 4311
Db 2281 TATGGGGTTTCAGATGATTAATTTGCTTCTCCGTTGGATATAAATGATATATTC 2340
Qy 4312 CAGGTAAAGGATATTCCTATTTTTCAGACAAAAAATACCAACACATGATGATCGAGGT 4371
Db 2341 CAGGTAAAGGATATTCCTATTTTTCAGACAAAAAATACCAACACATGATGATCGAGGT 2400
Qy 4372 GCAAGAGCATTTGCCCCATGATGCAAGAGGTCTTCCATCTGGAACTGGAGGTCTTTGTAAAA 4431
Db 2401 GCAAGAGCATTTGCCCCATGATGCAAGAGGTCTTCCATCTGGAACTGGAGGTCTTTGTAAAA 2460
Qy 4432 AATTCCTTTTCACTTGTCTACGACAGCAGATGAGTCTTACGGAATTAATGAATTCATCCAT 4491
Db 2461 AATTCCTTTTCACTTGTCTACGACAGCAGATGAGTCTTACGGAATTAATGAATTCATCCAT 2520
Qy 4492 TCAGATCCCTCTCTGTTTTTGTAGAAAGTACAGACACTGGACTACAGGAACATACAGAT 4551
Db 2521 TCAGATCCCTCTCTGTTTTTGTAGAAAGTACAGACACTGGACTACAGGAACATACAGAT 2580
Qy 4552 GATAACTGCCCTTTATTTGTGTCTGTTATTTGAAATTTCTGGGTTTTCCAGCCCGCAACCACTG 4611
Db 2581 GATAACTGCCCTTTATTTGTGTCTGTTATTTGAAATTTCTGGGTTTTCCAGCCCGCAACCACTG 2640
Qy 4612 AGTGAATATGATGATCATTCAGACTTTCAAGATATTCATATTTCTGATTTGTTGAGAGAG 4671
Db 2641 AGTGAATATGATGATCATTCAGACTTTCAAGATATTCATATTTCTGATTTGTTGAGAGAG 2700
Qy 4672 ACTATCCATTAATCTTTAGAAAGTGGTTCCCTCTAAAGTTTTTCGGGGATTTCTCGAATGAGT 4731
Db 2701 ACTATCCATTAATCTTTAGAAAGTGGTTCCCTCTAAAGTTTTTCGGGGATTTCTCGAATGAGT 2760
Qy 4732 GATGGGGTGTCTCAAGAGGCTCAGCTAGCAGCAACCAAGCAAGCAAGCAAGCAAGCAAGT 4791
Db 2761 GATGGGGTGTCTCAAGAGGCTCAGCTAGCAGCAACCAAGCAAGCAAGCAAGCAAGT 2820
Qy 4792 GTTAAACCAATTCAGATGATACCAATGTGCGGTATCTCTTCCGAAAGAAAGTTCTA 4851
Db 2821 GTTAAACCAATTCAGATGATACCAATGTGCGGTATCTCTTCCGAAAGAAAGTTCTA 2880
Qy 4852 AGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAACTAAATGTCAATGAGCTGGGCTTTTA 4911
Db 2881 AGATTAGTCAATTAATTTGAGTAGTTCAGTTTCAACTAAATGTCAATGAGCTGGGCTTTTA 2940
Qy 4912 ACAATTAAGGAGAGATTCCTCAAAATTTGATGACATATGCGCTTTTACTCTGAGGTTTCC 4971
Db 2941 ACAATTAAGGAGAGATTCCTCAAAATTTGATGACATATGCGCTTTTACTCTGAGGTTTCC 3000
Qy 4972 CATTTGCTGTCACTGTGCAATTCAGACTTCCGTTGCGAGGTTTCATACAGAAATTAATTT 5031
Db 3001 CATTTGCTGTCACTGTGCAATTCAGACTTCCGTTGCGAGGTTTCATACAGAAATTAATTT 3060
Qy 5032 CAAGATGCTACAGTTTCTACAAATGCAATGAGAGAGAGGCTGTGTTGGCAACACCA 5091
Db 3061 CAAGATGCTACAGTTTCTACAAATGCAATGAGAGAGAGGCTGTGTTGGCAACACCA 3120
Qy 5092 AAGCAACCTATAGTTGATACATCTGCTGAATCTCTGA 5127
Db 3121 AAGGACCTATAGTTGATACATCTGCTGAATCTCTGA 3156

LOCUS	AK173326	6816 bp	mRNA	linear	ROD 28-JUL-2004
DEFINITION	Mus musculus mRNA for mKIAA1999 protein.				
ACCESSION	AK173326				
VERSION	AK173326.1 GI:50511236				
KEYWORDS	FLI_CDNA.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1. Okazaki, N., Kikuno, R.F., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Saga, Y., Seino, S., Nishimura, M., Kaisho, T., Hoshino, K., Kitamura, H., Nagase, T., Ohara, O. and Koga, H.				
AUTHORS	Prediction of the Coding Nucleotide Sequences of Mouse Homologues of KIAA				
TITLE	Gene: IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences of cDNA Clones Randomly Sampled from Size-Fractionated Libraries				
JOURNAL	DNA Res. 11, 205-218 (2004)				
REFERENCE	2 (bases 1 to 6816)				
AUTHORS	Okazaki, N., Kikuno, R.F., Nagase, T., Ohara, O. and Koga, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-MAY-2004) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918) The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.				
FEATURES	Location/Qualifiers				
source	1. .6816				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/db_xref="taxon:10090"				
	/clone="mc903906"				
	/tissue_type="adult thymus"				
	/notes="vector:modified pBC SK+"				
gene	1. .6816				
	/genes="mKIAA1999"				
CDS	<1..2654				
	/genes="mKIAA1999"				
	/notes="CDS is predicted by in silico analysis. Start codon is not identified."				
	/codon_start=3				
	/evidence=not experimental				
	/product="mKIAA1999 protein"				
	/protein_id="BAD32604.1"				
	/db_xref="GI:50511237"				
	/translation="KOLEKWHKEYNSKYVDLIEEQLNEALTYRKPIDGDNVRRSNQ RLQRPHVLPVHLYGOLVHHKTCCHLLEVSIIITELCHNVRTPDLDKWDIKKLKASL WALGNIGSSNWGLNLQENVIPIIDLAKOCEVLISIRGTCVVVLGLIAKTKOGCDIL KCHSWDSVRSRKHLPVVPDDVEQLCNELSSVPSTLSNSESSTSRHNSSESAPSS MFPLEDRGSGSTSTFFLDINEDAPFYDRPGIKOKNSPFFGSSKLVRNRLIS LTLPTKKHRSDDPGKGLSENKSNKSNRIRLTLPESVDLNSHEDFTSSAQKSLQLE PSFVGNKHLIDAGSTPSIGENDLKFPSGTETHTRENTSRLVWVGSASSHIKIRSQ SFNPTDTTSGISMSSSPKEKTVAVDPTAMDTCGLSTVWSTKTVKTSHYLTPOSNH LSLKSNVSLVPKSGSHLPRAQSLKPSATIKSLADCNFSYTSRRDAFGYATLK RLQOORMPHLSHSEALAPKDVLTPTDITMKANSFESRLTPSRPMKALSVASLDKE DLLSPINHLTORSSVRNSVSATYGGSDDIIGLALPVDINDIIOIKDVPVPSKHV PPDDRGARMFSDHAGLSGAGGLVKNLSFHLRQOMSLEIMNSVHSDASLPLESTE DTGLQEDHNDCLYCVSIELLGFQPSNQLSSIICRSHDLQDIIPYSDWCEQTHNPLEV PKGFSIGSCSQASQEGESASTKSTELLGLVKITPDDTPMCRILLRKEVLRVLVNL SSVSTPKCHETGLLTKEKYPOTFDDICLYSEVSHLSHCTFRLQCRRIQLFQDVQ FLQMHSEAEAVLAIPQIVDSEAS"				
ORIGIN	Query Match 39.7%; Score 2037.8; DB 10; Length 6816; Best Local Similarity 86.0%; Pred. No. 0; Matches 2284; Conservative 0; Mismatches 367; Indels 6; Gaps 2;				
Qy	2474	CAAAACAATTGGAAAGTGGCACAGGAATACAACTCCAAATATGTTGACTGATTGAGG	2533		
Db	1	CAAAACAGTTGGAAAGTGGCACAGGAATATAAATTCAAAGTATGTTGACTGATTGAGG	60		

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-3590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 41 Row: b Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source

```

1..2430
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="IMAGE:4012732"
/tissue_type="Mammary tumor metastasized to lung."
MMTV-LTR/Mnt1 model. Expression driven by an MMTV-LTR
enhancer."
/clone_lib="NCI CGAP_Lu30"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
<1..2430
/gene="4921505C17Rik"
/db_xref="LocusID:78757"
<1..2391
/gene="4921505C17Rik"
/codon_start=1
/product="4921505C17Rik protein"
/protein_id="AAH22906.1"
/db_xref="GI:18605769"
/db_xref="LocusID:78757"
/db_xref="MGI:1926007"
/translation="QALRLVRKMITVNASLFPSSVANSLLIAGNDGLQERDMVRACI
AIIICLALQNPVWALRGGLNTILKNVIDCOLSRINEALITLHLLNHPKTRQYVRA
DVELERILAPYDHYRSPDTEAGOLKEDREARFLASKMGIIATFRSWAGIILNCKP
GNSGQSLIGVLICINMEIRRGLEVLVDIPLPLPVDTDFIEALLSVDPGRFQDSW
RUSDGFVAEAKTILPHRARSPDLMDNYLALILAFIRNGLEGLVEVITNSDDHIS
VRATILGLGELHMANITLPHSHSHLCLPTLMNNAASPDIPKELRLASALNCLR
FHEMKRGPKPSLHLDHIQKAIATHHRDQYLRVQKDI FVLKDTESALLINRDSQ
VLOHKNLWDMNDLITGLIKWPNVLRNYKQDLRFVRLIYFKPSPSKLYASDLD
LAKSKQLTVGQCFTEFLSESDGQYLEDVLDIOMLNASSGVKPSKPSLQNNGLL
TTLSQRYFLPGLTSCPHGVKLEKCSVFQCLNLCSLKNQDHLKLTVSSLDYSRD
GLARVLTLKIDACDKYLRVLLRNVEFFNNWGLLELLVQLHDKNKTIS
EALDILDEACEKAMHLIOMKPSALSHGLDGLLLRFLPSIPKGFSLNRYGVAK
OLEKHWEKNSYKDIIEQLNEALTYYRKPIDGDNVYRNSRQRLQRPVLPVLYLG
QLVHHTGTGCHLLEVSQSIITELCHNVRTPLDLKWDIKKLKSLMALVCSKQFSFQ"

```

ORIGIN

Query Match 38.9%; Score 1992.6; DB 10; Length 2430;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 2130; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 433 CAAGCACTTCGATCTAGTCAAGAAGATGATTCTGTAAGTCTTCCCTGTTTCTAGTTCT 492
Db 1 CAAGCCCTTCGCTTAGTCAGAAAGATGATTCTGTAAGTCTTCCCTGTTTCTAGTTCT 60
QY 493 GTGACCAACTCATTAATTCGAGTTGGAAATGATGGACTTCAAGAAAGACAGCAATGGTC 552

```

61 GTGGCCAACTCATTAATTCGGTTGGAAATGATGGCTTCAGGAACGAGACAGAAATGGTC 120
553 CGAGCATGCTATTCGCTATTCGTGAACTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTT 612
121 CGGGCTTCATCGCGCATCATCTGTGAATAGCACTTCAGAAACCCAGAGGTGGTGGCCCTT 180
613 CGAGGAGGACTAAACACCATATTTGAAAATGTGATTTGATTTGCCAATTAAGTCGAATAAAT 672
181 CGAGGAGGACTAAACACCATATTCGAAAATGTGATTTGATTTGCCAATTAAGTCGAATAAAT 240
673 GAGGCCCTTAATTAATTAATTTTGCACCTTTCTTAATTAATTAATTAATTAATTAATTAATTA 732
241 GAGGCTCTCATTAATTAATTTTGCACCTTTCTTAATTAATTAATTAATTAATTAATTAATTA 300
733 CGAGCTGATGTAGAAATTAAGAGAGAAATTTAGCACCTTATCTGATTTCCATCTACAGACAT 792
301 CGAGCTGATGTAGAAATTAAGAGAGAAATTTAGCACCTTATCTGATTTCCATCTACAGACAT 360
793 AGTCCAGATACAGCTGAAGGACAGCTCAAGAAAGACAGAGACAGACAGATTTCTAGCCAGT 852
361 AGTCCAGATACAGCTGAAGGACAGCTCAAGAAAGACAGAGACAGACAGATTTCTAGCCAGT 420
853 AAAATGGGAATCATAGCAACATTCGATCATGGGAGGATTAATTAATTAATTAATTAATTAATTA 912
421 AAGATGGGAATCATTTGCGACATTCGCTCGTGGGAGGATTAATTAATTAATTAATTAATTAAT 480
913 GGAATTTCTGGGCTCAGTCTTAATAGGAGTACTTTTCATACCAAAATATGGAATAAGG 972
481 GGAATTTCTGGGATACAGTCTTAATAGGAGTACTTTTCATACCAAAATATGGAATAAGG 540
973 CGAGTCTACTTGAAGTCTTTATGATATATTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1032
541 CGAGTCTCTGCAAGTCTCTTAAGACATCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
1033 GAGTTTCATAGAAGCACTACTCAGTGTAGATCAGGAGAGGTTCGAAAGACAGTGTGGAGGCTT 1092
601 GAGTTTCATAGAAGCCCTCTCAGTGTAGATCAGGAGAGATTCGAGGACAGTGTGGAGGCTT 660
1093 TCAGATGCTTTGTGGGAGCTGAGGCAAAACTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1152
661 TCCGATGGCTTTGTGGGAGCTGAGGCAAAACTATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
1153 CCAGACCTCATGGATAATTAATTTGGCACTGATACCTCTCTGCAATTTATTCGTAATGGACTT 1212
721 CCAGACCTCATGGCAATTAATTTGGCTCTGATACCTCTCTGCAATTTATTCGTAATGGACTT 780
1213 TTAGAGGTCTAGTTGAAGTGAATAACAAACAGTGTATGATCATCTCAGTTAGAGCTTACC 1272
781 TTAGAGGTCTAGTTGAAGTGAATAACAAACAGTGTATGATCATCTCAGTTAGAGGTGACC 840
1273 ATCCCTTTTAGGAGAGCTTTTACATATGCAACAACTTTCTCTCTCTCTCTCTCTCTCTCTCT 1332
841 ATCCCTTTTAGGAGAGCTCTCTGATATGGCAACCACTCTCTCTCTCTCTCTCTCTCTCTCT 900
1333 CATTTTACACTGCTTCCCAACCTTAATTAATTAATTTGCTGATCTCTCTGATATCCCAAGGAA 1392
901 CATCTACACTGCTTCCCAACCTCTATGAACATGGCTGGCTGCTCTCTGATATTTCCCAAGGAA 960
1393 AAGAGCTCGGAGCGAGCTGAGCCCTTGAATCTGTTTAAACCGTTCCTGATGAATGAAGAA 1452
961 AAGAGCTCGGAGCGAGCTGAGCCCTTGAATCTGTTTAAACCGTTCCTGATGAATGAAGAA 1020
1453 CGAGGACCTTAAGCTTATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1512
1021 CGAGGACCTTAAGCTTATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080
1513 CACCAAGAAACGGGATCAGTATCTCCGAGTTCAAGAAAGATATATTTATCTTCTTCTTCTTCT 1572
1081 CACCAAGAAACGGGATCAGTATCTCCGAGTTCAAGAAAGATATATTTATCTTCTTCTTCTTCT 1140
1573 GAGGAGCTCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 1632

```

Db 1141 GAGGAAGCTCTTTTAAATAAACCCTTAGAGACGCCAAGTCCTTTCAGACATAAAGAGAAATCTT 1200
Qy 1633 GAATGGAATTTGGAATCTTATAGGACCAATTTCTTAAGTGGCCAAAATGTAAATCTAAGAAAC 1692
Db 1201 GACTGGGATTTGGAATCTGATTGGGACCAATCTTTAAGTGGCCAAAATGTAAATCTAAGAAAC 1260
Qy 1693 TATAAAGATGAACAGTTTACACAGGTTTGTGACGAGCTACTTATTTTATTAAGCCGACG 1752
Db 1261 TATAAAGATGAGCAGTTGACAGGTTTGTGCGCAGACTTCTTTACTTTTACAAGCCGACG 1320
Qy 1753 AGTAAATATATGCCAACTGATCTGGATTTTGCACAGGCCAAGCTCAAGCTCAAGCTTCTA 1812
Db 1321 AGCAAACTGTAGCTAGTCTGGATCTGGACTGGCCAAAGTCAAGCAGCTCACAGTTGTC 1380
Qy 1813 GGTTCGCCAGTTTACAGAAATTTCTTCTGAATCTGAAGAGGATGGCAAGGCTACTTAGAA 1872
Db 1381 GGTTCGAGTTTACAGAAATTTCTGCTCGAGTCTGAAGAGGATGGCAAGGATACTTAGAA 1440
Qy 1873 GATCTAGTAAAGGATATTTGTTCACTGGCTCAATGCTTCACTCTGGAATGAACCCGAAAGA 1932
Db 1441 GATCTCGTGAAGATATTTGTTCACTGGCTCAATGCTTCACTCTGGAATGAACCCGAGAGA 1500
Qy 1933 AGTCTTCAAAATATGTTTATTTGACCAACCCTTAGTCAACACTACTTTTATTTATTTGA 1992
Db 1501 AGCCTTCAAGAAATGTTTACTGACCAACCCTTAGTCAACACTACTTCTTATTTATCGGA 1560
Qy 1993 ACATCTTCTTGCCACCCCTCATGGAGTTAAATGCTGGAAAAATGCAAGTGTATTTCAAGTGT 2052
Db 1561 ACATCTCTTGTCTCCTCATGGAGTCAAAATGCTGGAAAAATGCAAGTGTCTTTCAGTGT 1620
Qy 2053 CTCCTTAATCTTGTCTTGGAAAAACAAGATCACTTGTCTTAAACTTTTACTTGTCTAGC 2112
Db 1621 CTCCTTAATCTTGTCTTGGAAAAACAAGATCACTTGTCTTAAACTTTTACTTGTCTAGC 1680
Qy 2113 TTGGACTATAGCAGATGGAATGGCTAGAGTCACTTCTTCAAAATTTTAACTGCGAGCT 2172
Db 1681 CTAGACTATAGCAGATGGAATGGCTAGAGTCACTTCTTCAAAATTTTAACTGCGAGCT 1740
Qy 2173 ACTGATGCTCAGAGCTCTGATATCTTCAACAAAAATTTAAGGTTATTTAGAGAGCTAAAGTT 2232
Db 1741 ACAGATGCTCAGAGCTCTGATATCTTCAACAAAAATTTAAGGTTATTTAGAGAGCTAAAGTT 1800
Qy 2233 GAATTTCTTAAATTTGGGAAATTTGAGTTGTTAGTGACCCAGCTACATGATATAAAAAA 2292
Db 1801 GAATTTCTTCAAAATTTGGGAAATTTGAGTTGTTAGTGACCTACATGATATAAAAAA 1860
Qy 2293 ACAGATTTCTTCAAGCTCTTGTATATCTTCAAGCAAAAAATTTAAGGTTATTTAGAGAGCTAAAGTT 2352
Db 1861 ACAATTTCTTCAAGCTCTGATATTTCTTGAAGAGCTTGTGAAGCAAGGCTTCAATCTT 1920
Qy 2353 CATGCTCTTATTCAGATGAACCCAGGCTTATCCACCTTGGAGACAAAGGTTTGTCTTCTC 2412
Db 1921 CATGCTCTTATTCAGATGAACCCAGGCTTATCCACCTTGGAGACAAAGGTTTGTCTTCTC 1980
Qy 2413 CTGCTGAGATTTCTTCCATTTCCAAAGGATTTTCTTATCTGAATGAAGAGGTTATGTA 2472
Db 1981 CTGCTGAGATTTCTTCCATTTCCAAAGGATTTTCTTATCTGAATGAAGAGGTTATGTT 2040
Qy 2473 GCAAAACAAATTTGGAAGTGGCAAGGAAATACAACTCCCAAATATGTTGATCTGATGAG 2532
Db 2041 GCAAAACAAATTTGGAAGTGGCAAGGAAATATAAATTCAAAGTATGTTGATCTGATGAG 2100
Qy 2533 GAACAACTCAATGAAGCACTTACTTACTTACCGGAGCTGTTGATGTTGATGAATATGTT 2592
Db 2101 GAACAACTCAATGAAGCTTCTACTTACATACAGGAGCCCTATTTGATGTTGATGAATATGTT 2160
Qy 2593 CGTCGAGTAAACCAAGATTTACAGGCTCTCAAGCTTACCTGCTTATACACTTATGGA 2652
Db 2161 CGTCGAGTAAACCAAGATTTACAGGCTCTCAAGCTTATGATGTTGATGAATATGTT 2220
Qy 2653 CAACCTAGTACCAATAAAACAGGCTGCAATTTGTTGGAAGTACAGAAATATTTATACAGAA 2712
Db 2221 CAACCTAGTACCAATAAAACAGGCTGCAATTTTACTGGAAGTACAGAGTATCATCACAGAG 2280

Qy 2713 CTCTGCTCGTAATTTGCTGTAACACGATTTGGATAAGTGGGAAGAAATTTAAAAAATCTGAAA 2772
Db 2281 CTCTGCTCAATTTGCTGACTTCCGACTTGGACAATGGGAAGACATTTAAGAGCTGAAA 2340
Qy 2773 GCACTCTTTTGGGCTTGG 2791
Db 2341 GCCTCTCTTTGGGCTTGG 2359
RESULT 12
LOCUS CQ411997 4657 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19068 from Patent WO0170979.
ACCESSION CQ411997
VERSION CQ411997.1 GI:41319778
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lee, J. and Lillie, J.
TITLE Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 19068 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1.4657
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 34.6%; Score 1772.8; DB 6; Length 4657;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 2; Indels 97; Gaps 1;

Qy 3148 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCTTGATATCAATGAAGATACAGAG 3207
Db 4634 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCTTGATATCAATGAAGATACAGAG 4575
Qy 3208 CCAACATTTTATGACCCATCTGGACCCATTAAGAGTAAATTTCAATCCCTTTCTTGTCT 3267
Db 4574 CCAACATTTTATGACCCATCTGGACCCATTAAGAGTAAATTTCAATCCCTTTCTTGTCT 4515
Qy 3268 TCTAGTAACTTGTGAAGATTCGTATCTTAAATTCGCTTACTTTGCCCTTAAACAAACAT 3327
Db 4514 TCTAGTAACTTGTGAAGATTCGTATCTTAAATTCGCTTACTTTGCCCTTAAACAAACAT 4455
Qy 3328 CGTAGTACAGTGTATCCAAAGGAGGAAATTTATCATCTGAAGTAAAGACAGACAGG 3387
Db 4454 CGTAGTACAGTGTATCCAAAGGAGGAAATTTATCATCTGAAGTAAAGACAGACAGG 4395
Qy 3388 CGAATCAGAACACTTACGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACCC 3447
Db 4394 CGAATCAGAACACTTACGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACCC 4335
Qy 3448 ATATCCACTGTACAGAAAAATTAACAATTAGAGACTTTCATTATGGGAAATGACACAT 3507
Db 4334 ATATCCACTGTACAGAAAAATTAACAATTAGAGACTTTCATTATGGGAAATGACACAT 4275
Qy 3508 GAAGACACTGGTAGTACACCAAGCATTTGGAGAAATGACTTTAAATTTTACCAGAAATTTT 3567
Db 4274 GAAGACACTGGTAGTACACCAAGCATTTGGAGAAATGACTTTAAATTTTACCAGAAATTTT 4215
Qy 3568 GGTACAGAAATCAGACAGAAAAATTAACAAGCGAGAGGTTTAGTAGTAAAGTTCAACG 3627
Db 4214 GGTACAGAAATCAGACAGAAAAATTAACAAGCGAGAGGTTTAGTAGTAAAGTTCAACG 4155
Qy 3628 AGCTCAGATATGAAGATACGTTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 3687
Db 4154 AGCTCAGATATGAAGATACGTTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATA 4095


```
QY 3688 AGTTCAATGAGCTCAAGTCTCTCACGAGAGACAGTAGGTAGATGCTCAACACTATGGAC 3747
Db 4094 AGTTCAATGAGCTCAAGTCTCTCACGAGAGACAGTAGGTAGATGCTCAACACTATGGAC 4035
QY 3748 ACAGACTGTGGAAGACAGTAGTACTGTGGTAAGTACTAAACCTATTAAAGACAAGCCACTAT 3807
Db 4034 ACAGACTGTGGAAGACAGTAGTACTGTGGTAAGTACTAAACCTATTAAAGACAAGCCACTAT 4009
QY 3808 TTGACGCCACAGCTTAACCAATCTGTCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCT 3867
Db 4008 ----- 4009
QY 3868 CCAGGTTCTTCTCATACGCTTCTTAGAGAGACAGTCCCTTAAAGCAACCTCTATTGCT 3927
Db 4008 ---GGTCTTCTCATAGGCTTCTTAGAGAGACAGTCCCTTAAAGCAACCTCTATTGCT 3952
QY 3928 ACAATTTAAAGTCTAGCAGATTCTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGC 3987
Db 3951 ACAATTTAAAGTCTAGCAGATTCTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGC 3992
QY 3988 TATGCTACTGTAAGAAAGACTACAGCAACAAAGAAATGCAATCTCTTCTCACTCTGAA 4047
Db 3891 TATGCTACTGTAAGAAAGACTACAGCAACAAAGAAATGCAATCTCTTCTCACTCTGAA 3832
QY 4048 GCTTTGGCATCTCCAGCAAGAGATGCTATTCTTACTGATACCATCAACATGAAGGCCAAC 4107
Db 3831 GCTTTGGCATCTCCAGCAAGAGATGCTATTCTTACTGATACCATCAACATGAAGGCCAAC 3772
QY 4108 AGTTTGGAGTCCAGATTAAACCAAGCAGGTTTCAAGAAAGCTTTAAGTTATGATCATTA 4167
Db 3771 AGTTTGGAGTCCAGATTAAACCAAGCAGGTTTCAAGAAAGCTTTAAGTTATGATCATTA 3712
QY 4168 GATAAAGAGATTATTGAGTCTCTATTATCAAAATACCTGCAACGATCTTCTCAGTG 4227
Db 3711 GATAAAGAGATTATTGAGTCTCTATTATCAAAATACCTGCAACGATCTTCTCAGTG 3652
QY 4228 CGGTCCATGCTGTCAGAGTCCACATATGGGGTTCAGATGATTACATTTGCTTGTCTTC 4287
Db 3651 CGGTCCATGCTGTCAGAGTCCACATATGGGGTTCAGATGATTACATTTGCTTGTCTTC 3592
QY 4288 CCGGTGGATATAATGATATATTCCAGGTAAAGGATATCCCTTATTTTCAGACAAAAAC 4347
Db 3591 CCGGTGGATATAATGATATATTCCAGGTAAAGGATATCCCTTATTTTCAGACAAAAAC 3532
QY 4348 ATACCAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4407
Db 3531 ATACCAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3472
QY 4408 TCTGGAACTGGAGTCTTGTAAATAATCTTTTCACTTGTGTCAGACAGAGATGATCTT 4467
Db 3471 TCTGGAACTGGAGTCTTGTAAATAATCTTTTCACTTGTGTCAGACAGAGATGATCTT 3412
QY 4468 ACCGAAATTAATGAATTAATCAATCCATTCAGATGCTCTCTGTTTTAGAAAGTACAGAAG 4527
Db 3411 ACCGAAATTAATGAATTAATCAATCCATTCAGATGCTCTCTGTTTTAGAAAGTACAGAAG 3352
QY 4528 ACTGGAATACAGAAACATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4587
Db 3351 ACTGGAATACAGAAACATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3292
QY 4588 GGTTCACAGCCAGCAACCAAGTGGCAATATGATGATGATGATGATGATGATGATGATGATGAT 4647
Db 3291 GGTTCACAGCCAGCAACCAAGTGGCAATATGATGATGATGATGATGATGATGATGATGATGAT 3232
QY 4648 CCATATTTCTGATTTGTTGAGCAGACTATCCATAATCTTTAGAAAGTGTTCCTCTAAG 4707
Db 3231 CCATATTTCTGATTTGTTGAGCAGACTATCCATAATCTTTAGAAAGTGTTCCTCTAAG 3172
QY 4708 TTTTCGGGGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4767
Db 3171 TTTTCGGGGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3112
```

```
QY 4768 AAAAGCACAGAATTGTTACTAGGTGTTAAAACAAATTCAGATGATACACCAATGTCCCT 4827
Db 3111 AAAAGCACAGAATTGTTACTAGGTGTTAAAACAAATTCAGATGATACACCAATGTCCCT 3052
QY 4828 ATACTCTTTCGCAAGAAAGTCTTAAGATTAGTCAATTAATTTGAGTAGTTCAAGTTCAACT 4887
Db 3051 ATACTCTTTCGCAAGAAAGTCTTAAGATTAGTCAATTAATTTGAGTAGTTCAAGTTCAACT 2992
QY 4888 AAATGTCATGAGACTGGGCTTTTAAACAATTAAGGAGAGATATCCTCAACATTTGATGAC 4947
Db 2991 AAATGTCATGAGACTGGGCTTTTAAACAATTAAGGAGAGATATCCTCAACATTTGATGAC 2932
QY 4948 ATATGCTTTTACTCTGAGGTTCCTCCATTTGCTGTACACTGCACATTCAGACTTCCTGT 5007
Db 2931 ATATGCTTTTACTCTGAGGTTCCTCCATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2872
QY 5008 CGAGGTTTCATACAAAGAAATTTTCAAGATGATACAGTTTCTACAAATGATGAAGAAGCA 5067
Db 2871 CGAGGTTTCATACAAAGAAATTTTCAAGATGATACAGTTTCTACAAATGATGAAGAAGCA 2812
QY 5068 GAGGCTGTGTTGGCAACACACCAAGCAACCTATAGTTGATACATCTCTGAATCTGA 5127
Db 2811 GAGGCTGTGTTGGCAACACCAAGCAACCTATAGTTGATACATCTCTGAATCTGA 2752

RESULT 13
LOCUS CQ488314/c 4601 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence: 20181 from Patent WO0160860.
ACCESSION CQ488314
VERSION CQ488314.1 GI:41453933
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their
JOURNAL use
Patent: WO 0160860-A 20181 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
    source
        1..4601
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
    Query Match 34.5%; Score 1768; DB 6; Length 4601;
    Best Local Similarity 94.8%; Pred. No. 0;
    Matches 1878; Conservative 0; Mismatches 5; Indels 97; Gaps 1;

QY 3148 GACCGTTTGGCAGCAGCTCTACTAGTACATTTTCTTGTATCATCAATGAAGTACAGAG 3207
Db 4580 GACCGTTTGGCAGCAGCTCTACTAGTACATTTTCTTGTATCATCAATGAAGTACAGAG 4521
QY 3208 CCAACATTTTATGACCGATCTGACCCATTAAGGATAAAAATTCATTCCTTTCTTGCT 3267
Db 4520 CCAACATTTTATGACCGATCTGACCCATTAAGGATAAAAATTCATTCCTTTCTTGCT 4461
QY 3268 TCTAGTAAACTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCTTACAAAAACAT 3327
Db 4460 TCTAGTAAACTTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCTTACAAAAACAT 4401
QY 3328 CGTAGTAGCAGTATCCAAAAGAGGAGAAATTTATCATCTGAAAGTAAAGCAAGCAACAG 3387
Db 4400 CGTAGTAGCAGTATCCAAAAGAGGAGAAATTTATCATCTGAAAGTAAAGCAAGCAACAG 4341
QY 3388 CGAATCAGAACACTTACGAGGCCAGTGTGATTTTAAATCATAGTATGATTTTACACC 3447
Db 4340 CGAATCAGAACACTTACGAGGCCAGTGTGATTTTAAATCATAGTATGATTTTACACC 4281
```



```
QY 3448 ATATCCACTGTACAGAAAACATTACAATTAGAGACTTCAATTATGGGGAATAAGCACATT 3507
Db 4280 ATATCCACTGTACAGAAAACATTACAATTAGAGACTTCAATTATGGGGAATAAGCACATT 4221
QY 3508 GAAGACACTGGTAGTACACCAAGCAATTCGAGAAAATGACTTAAATTCACCAAGATTTT 3567
Db 4220 GAAGACACTGGTAGTACACCAAGCAATTCGAGAAAATGACTTAAATTCACCAAGATTTT 4161
QY 3568 GGTACAGAGAAATCAAGAGAAAATAACAAGCCGAGAGAGGTTAGTAGTAAAGTTCAACG 3627
Db 4160 GGTACAGAGAAATCAAGAGAAAATAACAAGCCGAGAGAGGTTAGTAGTAAAGTTCAACG 4101
QY 3628 AGCTCACATATGAAGATAGCTAGCCAAAGTTTCAATACAGACATCAACAAGTGGGATA 3687
Db 4100 AGCTCACATATGAAGATAGCTAGCCAAAGTTTCAATACAGACATCAACAAGTGGGATA 4041
QY 3688 AGTTCAATGAGCTCAAGTCCCTTCAGAGAGACAGTAGGTGTAGTGTACAACTATGGAC 3747
Db 4040 AGTTCAATGAGCTCAAGTCCCTTCAGAGAGACAGTAGGTGTAGTGTACAACTATGGAC 3981
QY 3748 ACAGACTGTGGAAGCATGAGTACTGTGTGTAAGTACTTAAACTATTTAAGACAAAGCCACTAT 3807
Db 3980 ACAGACTGTGGAAGCATGAGTACTGTGTGTAAGTACTTAAACTATTTAAGACAAAGCCACTAT 3955
QY 3808 TTGAGGCCACAGCTCTAACCATCTGTCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCT 3867
Db 3954 ----- 3955
QY 3868 CCAGGTTCTTCTCATAGCTTCTTAGAAGACACAGTCCCTTAAAGCACCCCTCTATTGCT 3927
Db 3954 ---GGTCTTCCGTCAGGCTTCTTAGAAGACACAGTCCCTTAAAGCACCCCTCTATTGCT 3898
QY 3928 ACAATTTAAAGCTAGCAGATTTGTAACCTTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGC 3987
Db 3997 ACAATTTAAAGCTAGCAGATTTGTAACCTTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGC 3838
QY 3988 TATGCTACACTGAAAAGACTACAGCAACAAAGAAATGCAATCACTTATCTCATCTGAA 4047
Db 3837 TATGCTACACTGAAAAGACTACAGCAACAAAGAAATGCAATCACTTATCTCATCTGAA 3778
QY 4048 GCTTTGCACTCCAGCAAAAGATGTCTATTTACTGTATACATCAACATGAAGGCCAAC 4107
Db 3777 GCTTTGCACTCCAGCAAAAGATGTCTATTTACTGTATACATCAACATGAAGGCCAAC 3718
QY 4108 AGTTTGAAGTCCAGATTAACCAAGCAGGTTTCATGAAGGCTTAAAGTTATGCAATCATTA 4167
Db 3717 AGTTTGAAGTCCAGATTAACCAAGCAGGTTTCATGAAGGCTTAAAGTTATGCAATCATTA 3658
QY 4168 GATAAAGAGATTTATGAGTCCCTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTG 4227
Db 3657 GATAAAGAGATTTATGAGTCCCTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTG 3598
QY 4228 CGGTCCATGGTCCAGTCCACATATGGGGTTTCAGATGATTAATGCTTCTGCTCTC 4287
Db 3597 CGGTCCATGGTCCAGTCCACATATGGGGTTTCAGATGATTAATGCTTCTGCTCTC 3538
QY 4288 CCGGTGGATATAAATGATATATTCAGGTAAGGATATTCCTATTTTCAGACAAAAAC 4347
Db 3537 CCGGTGGATATAAATGATATATTCAGGTAAGGATATTCCTATTTTCAGACAAAAAC 3478
QY 4348 ATACCAACCATGATGATCGAGGTCGAAGACATTTGCCATGATGACGAGGCTTTCCA 4407
Db 3477 ATACCAACCATGATGATCGAGGTCGAAGACATTTGCCATGATGACGAGGCTTTCCA 3418
QY 4408 TCTGGAAGTGGAGTCTTGTAATAAATTTCTTTTCACTGCTACAGACGATGAGTCTT 4467
Db 3417 TCTGGAAGTGGAGTCTTGTAATAAATTTCTTTTCACTGCTACAGACGATGAGTCTT 3358
QY 4468 ACGGAATTAATGAATTCATCACTCAGTGCCTCTCTGTTTTTATAGAAAGTACAGAGAC 4527
Db 3357 ACGGAATTAATGAATTCATCACTCAGTGCCTCTCTGTTTTTATAGAAAGTACAGAGAC 3298
QY 4528 ACTGGACTACAGGAACATACAGATGATAAATCTGCCTTTATTGTGTCTGTATTGAAATTCGT 4587
```

```
Db 3297 ACTGGACTACAGGAACATACAGATGATAAATGCGCTTTATTGTGTGTATTGAAATTCGT 3238
QY 4588 GGTTCAGGCCACCAACCAACTGAGTGCATATGTAGTCAATTCAGACTTTCAAGATATT 4647
Db 3237 GGTTCAGGCCACCAACCAACTGAGTGCATATGTAGTCAATTCAGACTTTCAAGATATT 3178
QY 4648 CCATATTCTGATTTGGTGTGAGCAGACTATCCATATCTTTAGAGTGGTTCCCTCTAAG 4707
Db 3177 CCATATTCTGATTTGGTGTGAGCAGACTATCCATATCTTTAGAGTGGTTCCCTCTAAG 3118
QY 4708 TTTTCGGGGATTTCTGATGCGAGTATGGGGTGTCTCAAGAGGCTAGCTAGCAGCACC 4767
Db 3117 TTTTCGGGGATTTCTGATGCGAGTATGGGGTGTCTCAAGAGGCTAGCTAGCAGCACC 3058
QY 4768 AAAAGCACAGAAATTTTACTAGGTGTTAAACAATTCAGATGATACACCAATGTGCCGT 4827
Db 3057 AAAAGCACAGAAATTTTACTAGGTGTTAAACAATTCAGATGATACACCAATGTGCCGT 2998
QY 4828 ATACTCTCTCGCAAGAGTTTCTAAGATTAGTCAATTTAGTAGTTCAGTTTCAACT 4887
Db 2997 ATACTCTCTCGCAAGAGTTTCTAAGATTAGTCAATTTAGTAGTTCAGTTTCAACT 2938
QY 4888 AAATGTCATGAGACTGGCTTTTAAACAATTAAGAGAAATGATCTCTCAACAATTTGATGAC 4947
Db 2937 AAATGTCATGAGACTGGCTTTTAAACAATTAAGAGAAATGATCTCTCAACAATTTGATGAC 2878
QY 4948 ATATGCTTACTCTGAGGTTTCCCATTTGCTGTGCACCTGCACATTCAGACTTCCCGT 5007
Db 2877 ATATGCTTACTCTGAGGTTTCCCATTTGCTGTGCACCTGCACATTCAGACTTCCCGT 2818
QY 5008 CGGAGGTTTCATACAAAGAAATTTTCAAGATGTACAGTTTCTCAAAATGCATGAAGAACA 5067
Db 2817 CGGAGGTTTCATACAAAGAAATTTTCAAGATGTACAGTTTCTCAAAATGCATGAAGAACA 2758
QY 5068 GAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTAGTATACATCTGCTGAATCTGA 5127
Db 2757 GAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTAGTATACATCTGCTGAATCTGA 2698

RESULT 14
CQ494141/c
LOCUS CQ494141 4602 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 26008 from Patent WO0160860.
ACCESSION CQ494141
VERSION CQ494141.1 GI:41459760
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 26008 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source Location/Qualifiers
1..4602
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 34.5%; Score 1768; DB 6; Length 4602;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 5; Indels 97; Gaps 1;

QY 3148 GACCGGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTGATATCAATGAAGATACAGAG 3207
Db 4580 GACCGGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTGATATCAATGAAGATACAGAG 4521
QY 3208 CCAACATTTTATGACCGATCTGGACCCCAAGGATAAAATTCATTTCTTCTTCTTCT 3267
```

Db	4520		CCACATTTTATGACCCGATCTGGACCCATAAAGATATAAAATTCATTCCCTCTCTTTGCT	4461
Qy	3268		TCTAGTAAACTTGTGAAGATCGTATCTTAAATTCGCTTACCTTTGCCTTAAACAAAAAAACAT	3327
Db	4460		TCTAGTAAACTTGTGAAGATCGTATCTTAAATTCGCTTACCTTTGCCTTAAACAAAAAAACAT	4401
Qy	3328		CGTAGTAGCAGTGATCCAAAAGGAGGGAAATATATCATCTGAAAGTAAAGACAAGCAACAGG	3387
Db	4400		CGTAGTAGCAGTGATCCAAAAGGAGGGAAATATATCATCTGAAAGTAAAGACAAGCAACAGG	4341
Qy	3388		CGAATCAGAACACATTACGGAGCCAGCTGTGATTTTAATCATAGTGATGATTTTACACCC	3447
Db	4340		CGAATCAGAACACATTACGGAGCCAGCTGTGATTTTAATCATAGTGATGATTTTACACCC	4281
Qy	3448		ATATCCACTGTPACAGAAACATTAACAATTAGAGACTTCATTATATGGGGAATAAGCACATTT	3507
Db	4280		ATATCCACTGTPACAGAAACATTAACAATTAGAGACTTCATTATATGGGGAATAAGCACATTT	4221
Qy	3508		GAAGACACTGGTAGTACACCAAGCAATTTGGAGAAAATGACTTTAAATTCACCAAGAAATTTT	3567
Db	4220		GAAGACACTGGTAGTACACCAAGCAATTTGGAGAAAATGACTTTAAATTCACCAAGAAATTTT	4161
Qy	3568		GGTACAGAGAAATCACAGAGAAAAATACAAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACG	3627
Db	4160		GGTACAGAGAAATCACAGAGAAAAATACAAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACG	4101
Qy	3628		AGCTCACATATGAAGATACGTAGGCCAAAGTTTCAATACAGACACTACAAACAAGTGGCATA	3687
Db	4100		AGCTCACATATGAAGATACGTAGGCCAAAGTTTCAATACAGACACTACAAACAAGTGGCATA	4041
Qy	3688		AGTTTCAATAGAGCTCAAGTCCCTTACAGAGACAGTAGGTGTAGATGCTACCACTATTGGAC	3747
Db	4040		AGTTTCAATAGAGCTCAAGTCCCTTACAGAGACAGTAGGTGTAGATGCTACCACTATTGGAC	3981
Qy	3748		ACAGACTGTGGAGCATGAGTACTGTGGTAAAGTACTTAAACTATTAAGACAAGCCACTAT	3807
Db	3980		ACAGACTGTGGAGCATGAGTACTGTGGTAAAGTACTTAAACTATTAAGACAAGCCACTAT	3955
Qy	3808		TTGAGGCCACAGTCTAACCATCTGTCTCTCTCCAAATCAAATTCGGTGTCCCTGGTGCCT	3867
Db	3954		TTGAGGCCACAGTCTAACCATCTGTCTCTCTCCAAATCAAATTCGGTGTCCCTGGTGCCT	3955
Qy	3868		CCAGGTTCTTCTCATACGCTTCTTAGAAGACACAGTCCCTTAAAGCACCCCTCTATTGCT	3927
Db	3954		CCAGGTTCTTCTCATACGCTTCTTAGAAGACACAGTCCCTTAAAGCACCCCTCTATTGCT	3954
Qy	3928		ACAATTTAAAGTCTTAGCAGATTTGACTTTAGTTTACAAAGTCTAGAGATGCTTTTGGC	3987
Db	3897		ACAATTTAAAGTCTTAGCAGATTTGACTTTAGTTTACAAAGTCTAGAGATGCTTTTGGC	3838
Qy	3988		TATGCTACACTGAAAAGACTACAGCAACAAGAAATGCATCCATCTTATCTCACTCTGAA	4047
Db	3837		TATGCTACACTGAAAAGACTACAGCAACAAGAAATGCATCCATCTTATCTCACTCTGAA	3778
Qy	4048		GCCTTGGCATCTCGAGCAAAAGATGTGCTATTTACTGTATACATACCATGAAGGCCAAC	4107
Db	3777		GCCTTGGCATCTCGAGCAAAAGATGTGCTATTTACTGTATACATACCATGAAGGCCAAC	3718
Qy	4108		AGTTTTGAGTCCAGATTAAACCAAGCAGGTTCCATGAAAGCCCTTAAGTTATGTCATCATTTA	4167
Db	3717		AGTTTTGAGTCCAGATTAAACCAAGCAGGTTCCATGAAAGCCCTTAAGTTATGTCATCATTTA	3658
Qy	4168		GATAAAGAGATTTATTTAGTCTCTATTAATCAAAATACCCCTGCAACGATCTTCTCAGTG	4227
Db	3657		GATAAAGAGATTTATTTAGTCTCTATTAATCAAAATACCCCTGCAACGATCTTCTCAGTG	3598
Qy	4228		CGGTCCATGTGTCCAGTGCCACATATGGGGTTCCAGATGATTTACATTCGCTCTGCTCTC	4287
Db	3597		CGGTCCATGTGTCCAGTGCCACATATGGGGTTCCAGATGATTTACATTCGCTCTGCTCTC	3538
Qy	4288		CCGGTGGATATAAATGATATATTCCAGGTAAGGATATTCCCTATTATTTCCAGCAAAAAAC	4347

Db	3537	CCGGTGGATATAAATGATATATATTTCCAGGTAAAGGATATTCCTCTATTTTTCAGACAAAAC	3478
Qy	4348	ATACCAACACATGATGATCGAGGTGCAAGAGCATTTGGCCCATGATGATGAGGAGTCTTCCA	4407
Db	3477	ATACCAACACATGATGATCGAGGTGCAAGAGCATTTGGCCCATGATGAGGAGTCTTCCA	3418
Qy	4408	TCTGGAACTGGAGGTCTTGTGTAATAAATTTCTTTTCACTTGTCTAGCAGCAGATGAGTCTT	4467
Db	3417	TCTGGAACTGGAGGTCTTGTGTAATAAATTTCTTTTCACTTGTCTAGCAGCAGATGAGTCTT	3358
Qy	4468	ACGGAAATATGAATTCAAATCCATTCAGATGCCCTCTCTGTCTTTTGAAGAGTACAGAAGAC	4527
Db	3357	ACGGAAATATGAATTCAAATCCATTCAGATGCCCTCTCTGTCTTTTGAAGAGTACAGAAGAC	3298
Qy	4528	ACTGGACTACAGGAACATACAGATGATAACTGGCTTTTATGTGCTGTGATTTGAAATTCG	4587
Db	3297	ACTGGACTACAGGAACATACAGATGATAACTGGCTTTTATGTGCTGTGATTTGAAATTCG	3238
Qy	4588	GGTTTCAGGCCAGCAACCACTGAGTGCATATGATGATCATTCAGACTTTTCAAGATATT	4647
Db	3237	GGTTTCAGGCCAGCAACCACTGAGTGCATATGATGATCATTCAGACTTTTCAAGATATT	3178
Qy	4648	CCATATTCTGATTGGTGTGAGCAGACTATCCATAATTCCTTTAGAACTGGTTCCTCTAAG	4707
Db	3177	CCATATTCTGATTGGTGTGAGCAGACTATCCATAATTCCTTTAGAACTGGTTCCTCTAAG	3118
Qy	4708	TTTTTCGGGGATTTCTGGATCAGTGATGGGGTGTCTCAAGAGGCTCAGCTAGCAGCAC	4767
Db	3117	TTTTTCGGGGATTTCTGGATCAGTGATGGGGTGTCTCAAGAGGCTCAGCTAGCAGCAC	3058
Qy	4768	AAAAGCACAGATTGTTTACTGAGTGTATAAAACAATTCAGATGATACCAATGTGCCGT	4827
Db	3057	AAAAGCACAGATTGTTTACTGAGTGTATAAAACAATTCAGATGATACCAATGTGCCGT	2998
Qy	4828	ATACTCTTCGCAAGAAAGTTCTAAGATTAGTCAATTAATTTGAGTAGTTTCACTTTCACT	4887
Db	2997	ATACTCTTCGCAAGAAAGTTCTAAGATTAGTCAATTAATTTGAGTAGTTTCACTTTCACT	2938
Qy	4888	AAATGTCATCAGACTGGGCTTTTAAACAATTAAGGAGAATATCTTCAAAACATTTGATGAC	4947
Db	2937	AAATGTCATCAGACTGGGCTTTTAAACAATTAAGGAGAATATCTTCAAAACATTTGATGAC	2878
Qy	4948	ATATGCCCTTTACTCTGAGGTTTCCCATTGCTGTCTACACTGCACATTCAGACTTCGGTGT	5007
Db	2877	ATATGCCCTTTACTCTGAGGTTTCCCATTGCTGTCTACACTGCACATTCAGACTTCGGTGT	2818
Qy	5008	CGAGGTTTCATACAGAAATATTTCGAAGATGTACAGTTCTTCAAAATGATGAAGACA	5067
Db	2817	CGAGGTTTCATACAGAAATATTTCGAAGATGTACAGTTCTTCAAAATGATGAAGACA	2758
Qy	5068	GAGGCTGTGTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTCAATTCCTGA	5127
Db	2757	GAGGCTGTGTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTCAATTCCTGA	2698
RESULT	15		
LOCUS	CQ726625	1649 bp	DNA
DEFINITION	Sequence 12559 from Patent WO02068579.		
ACCESSION	CQ726625		
VERSION	CQ726625.1	GI:42290268	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 12559 06-SEP-2002; PE Corporation (NY) (US)		

FEATURES		Location/Qualifiers
source		1..1649
		/organism="Homo sapiens"
		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"
ORIGIN		
Query Match		29.1%; Score 1494; DB 6; Length 1649;
Best Local Similarity		92.1%; Pred. No. 0;
Matches 1645; Conservative		0; Mismatches 0; Indels 141; Gaps 1;
Qy	3128	GTATGTTTCATATTCGAGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTG 3187
Db	1	GTATGTTTCATATTCGAGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTG 60
Qy	3188	ATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCCAATGAAGATATAA 3247
Db	61	ATATCAATGAAGATACAGAGCCAACTTTTATGACCGATCTGGACCCCAATGAAGATATAA 120
Qy	3248	ATTCATTTCCCTTTCTTTGCTTTCTAGTAACTTTGTGAAGAATCGTATCTTAAATTCGCTTA 3307
Db	121	ATTCATTTCCCTTTCTTTGCTTTCTAGTAACTTTGTGAAGAATCGTATCTTAAATTCGCTTA 180
Qy	3308	CTTTGCCCTAACAAAAACATCGTAGTAGAGTATCCAAAAGGAGGAAATATATCATCTG 3367
Db	181	CTTTGCCCTAACAAAAACATCGTAGTAGAGTATCCAAAAGGAGGAAATATATCATCTG 240
Qy	3368	AAAGTAAGACAAAGCAACAGGCGAATCAGACACTTTACGGAGCCAGGTGTTGATTTTAATC 3427
Db	241	AAAGTAAGACAAAGCAACAGGCGAATCAGACACTTTACGGAGCCAGGTGTTGATTTTAATC 300
Qy	3428	ATAGTGATGATTTTACACCCATATCCACTGTACAGAAAAACATTAACAATTAGAGACTTTCAT 3487
Db	301	ATAGTGATGATTTTACACCCATATCCACTGTACAGAAAAACATTAACAATTAGAGACTTTCAT 360
Qy	3488	TTATGGGGAATAAGCAATTAAGACACTGTGTAGTACACCAAGCAATTTGGAGAAATGACT 3547
Db	361	TTATGGGGAATAAGCAATTAAGACACTGTGTAGTACACCAAGCAATTTGGAGAAATGACT 420
Qy	3548	TAAATTTACCAAGAAATTTGGTACAGAGATACAGAGAAATACAGCCGAGAGAGGT 3607
Db	421	TAAATTTACCAAGAAATTTGGTACAGAGATACAGAGAAATACAGCCGAGAGAGGT 480
Qy	3608	TAGTAGTAGAAGTTTCAACGAGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAG 3667
Db	481	TAGTAGTAGAAGTTTCAACGAGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAG 540
Qy	3668	ACACTACAAAGTGGCATAGTTCAATGAGCTCAAGTCCCTTACAGAGACAGTAGGTG 3727
Db	541	ACACTACAAAGTGGCATAGTTCAATGAGCTCAAGTCCCTTACAGAGACAGTAGGTG 600
Qy	3728	TAGATGCTACAACTATGGACACAGCTGTGGAAGCATGAGTACTGTGTAAGTACTAAAA 3787
Db	601	TAGATGCTACAACTATGGACACAGCTGTGGAAGCATGAGTACTGTGTAAGTACTAAAA 660
Qy	3788	CTATTAAGACAAGCCACTATTTGACGCCACAGTCTAACCATCTCTCTCCAAATCAA 3847
Db	661	CTATTAAGACAAGCCACTATTTGACGCCACAGTCTAACCATCTCTCTCCAAATCAA 720
Qy	3848	ATTCGGTGTCCCTGGTCCCTCAGGTTCTTCTCATACGCTTCTTAGAAGACACAGTCCC 3907
Db	721	ATTCGGTGTCCCTGGTCCCTCAGGTTCTTCTCATACGCTTCTTAGAAGACACAGTCCC 780
Qy	3908	TTAAGACACCTCTATTGCTCAATTAAGTCTAGCAGATTTAACTTTAGTTACACAA 3967
Db	781	TTAAGACACCTCTATTGCTCAATTAAGTCTAGCAGATTTAACTTTAGTTACACAA 840
Qy	3968	GTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAGATGCATC 4027
Db	841	GTTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAGATGCATC 900
Qy	4028	CATCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATTTACTGATA 4087

Search completed: April 9, 2005, 05:41:23
Job time : 14476.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 11:22:35 ; Search time 1705.98 Seconds
(without alignments)
17790.649 Million cell updates/sec

Title: US-10-782-244-2

Perfect score: 5127

Sequence: 1 atggcgcgatcgccgcg.....atacatctgctgaatcctga 5127

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

100% processing. Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

```
Database : N_Geneseq_16Dec04:
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5127	100.0	5127	13	ADR44435	Adr44435 Human p20
2	5127	100.0	5662	13	ADR44434	Adr44434 Human p20
3	3829	74.7	3829	13	ADQ89851	Adq89851 Antagonis
4	3800	74.1	9232	10	ADE71191	Adel71191 Novel hum
5	2026.4	39.5	2659	10	ADC30081	Adc30081 Human nov
C 6	1772.8	34.6	4657	5	ADL45178	Adl45178 Human ova
C 7	1768	34.5	4601	5	ABV20190	Abv20190 Human pro
C 8	1768	34.5	4602	5	ABV26019	Abv26019 Human pro
9	967.8	18.9	1334	10	ADC49295	Adc49295 Polypepti
10	874.8	17.1	3129	13	ADR08103	Adr08103 Full leng
11	744	14.5	1292	5	AAS63246	Aas63246 DNA encod
12	612.8	12.0	1765	4	AAI96373	Aai96373 Human neu
13	537.8	10.5	1185	5	AAS65245	Aas65245 DNA encod
14	537.8	10.5	1185	5	AAS80497	Aas80497 DNA encod
15	537.8	10.5	1185	10	ADC32002	Adc32002 Human nov
16	465.2	9.1	570	5	ADI67269	Adi67269 Human ova
17	465.2	9.1	570	5	ADI73667	Adi73667 Human ova
18	459	9.0	473	3	AACT5367	Aac75367 Human ORF
C 19	385.6	7.5	452	5	ABV09387	Abv09387 Human pro
C 20	383.4	7.5	427	5	ABV39532	Abv39532 Human pro

C	21	383.4	7.5	427	5	ABV30563	Human pro
	22	370	7.2	401	3	AAC00944	Human sec
C	23	357	7.0	770	4	AAI96374	Human neu
C	24	340.8	6.6	1347	5	AAS72571	DNA encod
C	25	339.8	6.6	4726	13	ADR07157	Full leng
C	26	336.4	6.6	416	5	ABV00218	Human pro
	27	334	6.5	341	5	ADL38904	Human ova
C	28	321	6.3	450	3	ADF56995	Urogenita
	29	280.2	5.5	726	6	ABS76822	Frog embr
C	30	166.8	3.3	567	12	ACH75616	Human gen
	31	163.6	3.2	170	12	ACH89342	Human gen
C	32	133.8	2.6	5769	4	ABL13629	Drosophil
	33	133.8	2.6	5769	13	ADQ89657	Antagonis
C	34	81	1.6	9139	4	ABL13628	Drosophil
	35	77.4	1.5	289	5	ADL44180	Human ova
C	36	73	1.4	185	5	ADL63388	Human ova
	37	57	1.1	185	5	ADL37796	Human ova
C	38	57	1.1	185	5	ADI72657	Human ova
	39	49	1.0	2000	8	ADAY1938	Human ova
C	40	48.2	0.9	4590	5	AAH24065	Rice gene
	41	45.4	0.9	2000	8	ADAY1938	Yeast AOD
C	42	43.8	0.9	2427	5	AAS98862	Rice gene
	43	43.8	0.9	2427	5	AAS69554	DNA encod
C	44	41.8	0.8	37515	6	ABQ66998	DNA encod
	45	41.6	0.8	14041	4	AAH48024	Human ang
C	46	41.6	0.8	14041	4	AAH48024	Internal

ALIGNMENTS

RESULT 1

ADR44435

ID ADR44435 standard; cDNA; 5127 BP.

XX

AC ADR44435;

19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1

DT 18-NOV-2004 (first entry)
XX
DE Human p200 encoding cDNA SEQ ID NO:2.
XX
KW mTOR-associated protein; mTOR-AP; cytostatic; antidiabetic; gene therapy;
KW mTOR-AP modulation; cancer; tumour; diabetes; human; p200; gene; ss.
XX

XX PN WO2004074448-A2.

XX XX

PD 02-SEP-2004.

XX

PF 18-FEB-2004; 2004WO-US004821.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

PR 18-FEB-2003; 2003US-
YY

PA (WHEED) WHITEHEAD

XX Sabatini DM, Kim D, Sarbasov DD; ✓
XX WPI; 2004-635565/61.
XX N-PSDB; ADR44435.
XX
XX New mTOR-associated protein (mTOR-AP) polypeptide, useful for treating or
XX preventing a disorder that is responsive to mTOR-AP modulation, e.g.
XX cancer or diabetes
XX
XX

Cancer Of Diabetes.

XX XX

PS PS

Claim 1; SEQ ID NO 2; 91pp; English.

XX XX

The present invention describes an isolated mTOR-associated protein (mTOR-AP) polypeptide comprising a sequence that is 70% identical to the 1708

CC amino acid sequence of SEQ ID NO:3 (ADR44436, P1) or 70% identical to the
CC 327 amino acid sequence of SEQ ID NO:6 (ADR44439, P2), or that is encoded
CC by a nucleic acid that hybridizes under high stringency conditions to a
CC complement sequence of the 951271 nucleic acid sequence of SEQ ID NO:2
CC (ADR44435, S1) or the 981 nucleic acid sequence of SEQ ID NO:5 (ADR44438,
CC S2). Also described: (1) an isolated antibody, or its fragment, which is
CC specifically immunoreactive with an epitope of an amino acid sequence of
CC P1 or P2; (2) the isolated nucleic acid encoding the polypeptide, or its
CC complement; (3) a vector comprising a nucleic acid sequence encoding an
CC mTOR-AP polypeptide; (4) an isolated host cell comprising the nucleic
CC acid; (5) a method of producing an mTOR-AP polypeptide; (6) a method for
CC detecting the presence of an mTOR-AP polypeptide; (7) a kit for detecting
CC an mTOR-AP polypeptide comprising the antibody and a detectable label for
CC detecting the antibody; (8) a method for detecting the presence of the
CC nucleic acid in a sample; (9) a kit comprising the probe or primer
CC nucleic acid and instructions for use; (10) an isolated, purified or
CC recombinant complex comprising an mTOR polypeptide and an mTOR-AP; (11) a
CC method for identifying a compound which modulates activity of an mTOR-AP-
CC polypeptide; (12) a method of inhibiting aberrant activity of an mTOR-AP-
CC expressing cell; (13) a method of treating or preventing a disorder that
CC is responsive to mTOR-AP modulation, in a subject; and (14) a transgenic
CC mouse having germline and somatic cells comprising a chromosomally
CC incorporated transgene that disrupts the genomic mTOR-AP gene and
CC inhibits expression of the gene, where the disruption comprises insertion
CC of a selectable marker sequence resulting in the transgenic mouse
CC exhibiting increased susceptibility to the formation of tumors as
CC compared to the wild type mouse. mTOR-AP sequences cyclostatic and
CC antidiabetic activities, and can be used in gene therapy. The
CC polypeptides, polynucleotides, compounds and methods are useful for
CC treating or preventing a disorder that is responsive to mTOR-AP
CC modulation, e.g. cancer or diabetes. The compound, which modulates an
CC mTOR-AP activity or expression is useful for the manufacture of a
CC medicament for treating a disease affected by aberrant mTOR-AP activity
CC or expression. The present sequence encodes human p200, which is an mTOR-
CC AP used in the exemplification of the present invention.

XX
SQ Sequence 5127 BP; 1626 A; 1028 C; 1049 G; 1424 T; 0 U; 0 Other;

Query Match 100.0%; Score 5127; DB 13; Length 5127;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCGGCGATCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC	60
DB	1	ATGGCGGCGATCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGAATGAC	60
QY	61	AGCGGCGAGGAGACGTCCTCGGCTGGATCTGACCCGAGAACCTTCTGATTAACGAGAG	120
DB	61	AGCGGCGAGGAGACGTCCTCGGCTGGATCTGACCCGAGAACCTTCTGATTAACGAGAG	120
QY	121	ATTCTCCAAATGTGGCCAGATTGCGGAGTATCAATATGAGAAAGCTAGGCCATCTG	180
DB	121	ATTCTCCAAATGTGGCCAGATTGCGGAGTATCAATATGAGAAAGCTAGGCCATCTG	180
QY	181	AATACTTTAAGCTCTTTGTGATATTGGCCAGTGAAGAAACCTGGGCTTTTAC	240
DB	181	AATACTTTAAGCTCTTTGTGATATTGGCCAGTGAAGAAACCTGGGCTTTTAC	240
QY	241	TATGAGGATATCATAAATTTGTTGGGTTAGCTTTTAAATTAAGAACAAAGAGTGCGA	300
DB	241	TATGAGGATATCATAAATTTGTTGGGTTAGCTTTTAAATTAAGAACAAAGAGTGCGA	300
QY	301	GCAGCAGGCTACGAGCGCTCGATATCTCAGAGCTCCAGATTTCTCCAGAGGTG	360
DB	301	GCAGCAGGCTACGAGCGCTCGATATCTCAGAGCTCCAGATTTCTCCAGAGGTG	360
QY	361	CTAAATTTGAAGTGGACTATTTAATAGCTAGGTGCAATTCACATCAACAGAGCAACGAG	420
DB	361	CTAAATTTGAAGTGGACTATTTAATAGCTAGGTGCAATTCACATCAACAGAGCAACGAG	420
QY	421	GTAGAGGACACAGCCTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG	480
DB	421	GTAGAGGACACAGCCTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG	480

QY	481	TTTCTAGTTCTGTGACCAACTCAATTAATTTGCAGTTGGAATGATGACTTCAAGAAAGA	540
DB	481	TTTCTAGTTCTGTGACCAACTCAATTAATTTGCAGTTGGAATGATGACTTCAAGAAAGA	540
QY	541	GACAGAAATGTCGAGCATGCAATTTGCCATTTATCTGTGAATCTAGCACTTCAGAAATCCAGAG	600
DB	541	GACAGAAATGTCGAGCATGCAATTTGCCATTTATCTGTGAATCTAGCACTTCAGAAATCCAGAG	600
QY	601	GTGTGGCCCTTCGAGGAGGACTAAACACCATTTTGAATAATGTGATGTCGAATTA	660
DB	601	GTGTGGCCCTTCGAGGAGGACTAAACACCATTTTGAATAATGTGATGTCGAATTA	660
QY	661	AGTCGAATAATGAGGCCCTAAATTTACTACAAATTTTGACCTTCTTAATCATCCAAAGACT	720
DB	661	AGTCGAATAATGAGGCCCTAAATTTACTACAAATTTTGACCTTCTTAATCATCCAAAGACT	720
QY	721	CGGCAGTATGTGCGAGCTGATGTAGAATTTAGAGAAATTTTACACCTATATCTGATTTT	780
DB	721	CGGCAGTATGTGCGAGCTGATGTAGAATTTAGAGAAATTTTACACCTATATCTGATTTT	780
QY	781	CACATACAGATAGTCCAGATACAGCTGAGGACAGCTCAAGAGACAGAGACGACGA	840
DB	781	CACATACAGATAGTCCAGATACAGCTGAGGACAGCTCAAGAGACAGAGACGACGA	840
QY	841	TTTCTAGCCAGTAAATTTGGGAATCATAGCAACATTTCCGATCATGGCAGGATTTAAT	900
DB	841	TTTCTAGCCAGTAAATTTGGGAATCATAGCAACATTTCCGATCATGGCAGGATTTAAT	900
QY	901	TTATGTAAACCTGGAATTTCTGGGATCCAGTCTCTAATAGAGTACTTTTGGATACCAAT	960
DB	901	TTATGTAAACCTGGAATTTCTGGGATCCAGTCTCTAATAGAGTACTTTTGGATACCAAT	960
QY	961	ATGGAATAAGCGGAGCTTACTTTGAGTGCTTTTATGATATATTTCTGCTTCTCTACT	1020
DB	961	ATGGAATAAGCGGAGCTTACTTTGAGTGCTTTTATGATATATTTCTGCTTCTCTACT	1020
QY	1021	GTGTGACTGAGGAGTTCATAGAAGCACTACTCAGTGTAGATCCAGGAGGTTCAGAGAC	1080
DB	1021	GTGTGACTGAGGAGTTCATAGAAGCACTACTCAGTGTAGATCCAGGAGGTTCAGAGAC	1080
QY	1081	AGTTGGAGGCTTTGAGATGGCTTTGTGGCAGCTGAGGCAAAACCTATTTCTTCTCTAT	1140
DB	1081	AGTTGGAGGCTTTGAGATGGCTTTGTGGCAGCTGAGGCAAAACCTATTTCTTCTCTAT	1140
QY	1141	GCAGATCCAGGCGAGCTCATGATTAATTTATTTGGCAGCTGATCTCTGCAATTTAT	1200
DB	1141	GCAGATCCAGGCGAGCTCATGATTAATTTATTTGGCAGCTGATCTCTGCAATTTAT	1200
QY	1201	CGTAATGCACTTTTAGAGGGCTTAGTTGAAGTGATAACAAAACAGTGATGATCATATCTCA	1260
DB	1201	CGTAATGCACTTTTAGAGGGCTTAGTTGAAGTGATAACAAAACAGTGATGATCATATCTCA	1260
QY	1261	GTAGAGCTACATCTTTTAGGAGAGCTTTTACATATGGCAACACAAATTTCTTCTCTAT	1320
DB	1261	GTAGAGCTACATCTTTTAGGAGAGCTTTTACATATGGCAACACAAATTTCTTCTCTAT	1320
QY	1321	TCACATAGCCATCATTTTACACTGCTTGCACCCCTTAATGAATATGCTGCTCTTTGAT	1380
DB	1321	TCACATAGCCATCATTTTACACTGCTTGCACCCCTTAATGAATATGCTGCTCTTTGAT	1380
QY	1381	ATCCCCAAGGAAAGAGAGCTGCGAGCCAGTCAGCCTTTGAACCTTTTAAACGCTTCCAT	1440
DB	1381	ATCCCCAAGGAAAGAGAGCTGCGAGCCAGTCAGCCTTTGAACCTTTTAAACGCTTCCAT	1440
QY	1441	GAAATGAAGAACGAGGACCTTAAGCTTATAGTCTTCAATTTAGACCAATTTTCAAGAA	1500
DB	1441	GAAATGAAGAACGAGGACCTTAAGCTTATAGTCTTCAATTTAGACCAATTTTCAAGAA	1500
QY	1501	GCAATTTGCAACACACAGAAACGGGATCAGTATCTCCGAGTTCCAGAAAGATATTTATC	1560
DB	1501	GCAATTTGCAACACACAGAAACGGGATCAGTATCTCCGAGTTCCAGAAAGATATTTATC	1560

3721 GTAGGTGTAGTGTACAACTATGACACAGACTGTGGAAGCATGAGTACTGTGTAAGT 3780
3781 ACTAAAACTATTAAAGAACGACCACTATTATTGAGCCACAGCTCTAACATCTCTCTCTCC 3840
3781 ACTAAAACTATTAAAGAACGACCACTATTATTGAGCCACAGCTCTAACATCTCTCTCTCC 3840
3841 AATCAAAATTCGGTGTCCCTGTGCTCCAGGTCTCTCTCATACGCTTCTCTAGAGAGCA 3900
3841 AATCAAAATTCGGTGTCCCTGTGCTCCAGGTCTCTCTCATACGCTTCTCTAGAGAGCA 3900
3901 CAGTCCCTTAAAGCACCCCTCTATTGTCTCAATTAAGTCTAGCAGATGTAACCTTTAGT 3960
3901 CAGTCCCTTAAAGCACCCCTCTATTGTCTCAATTAAGTCTAGCAGATGTAACCTTTAGT 3960
3961 TACAAAGTCTAGAGATGCTTTGGCTATGCTACACTCAAAAGACTACAGCAACAAAGA 4020
3961 TACAAAGTCTAGAGATGCTTTGGCTATGCTACACTCAAAAGACTACAGCAACAAAGA 4020
4021 ATGCATCCATCCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATT 4080
4021 ATGCATCCATCCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATT 4080
4081 ACTGATACCATCATGAAAGGCAAGCTTTTGGTCCAGTTCACATCAACCAAGCAGGTC 4140
4081 ACTGATACCATCATGAAAGGCAAGCTTTTGGTCCAGTTCACATCAACCAAGCAGGTC 4140
4141 ATGAAAGCCTTAAGTATGATCATATAGATAAAGAGATTTATTAGTCTCTATTATCAA 4200
4141 ATGAAAGCCTTAAGTATGATCATATAGATAAAGAGATTTATTAGTCTCTATTATCAA 4200
4201 AATACCTGCAACGATCTTCTGAGTCCGCTCCATGCTGCTCCAGTCCACATATGGGGT 4260
4201 AATACCTGCAACGATCTTCTGAGTCCGCTCCATGCTGCTCCAGTCCACATATGGGGT 4260
4261 TCAGATGATTAATGATGCTTCTGCTCTCCGCTGATATAAATGATATATTCAGGTAAG 4320
4261 TCAGATGATTAATGATGCTTCTGCTCTCCGCTGATATAAATGATATATTCAGGTAAG 4320
4321 GATATTCCTTATTTTACAGAAAAAATACACACATGATGATGATGATGATGATGATGAT 4380
4321 GATATTCCTTATTTTACAGAAAAAATACACACATGATGATGATGATGATGATGATGAT 4380
4381 TTTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4440
4381 TTTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4440
4441 CACTTGTCTACGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500
4441 CACTTGTCTACGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500
4501 TCTCTGTTTAAAGTACAGACACTGAGTCTTACGGAATTAATGAATTCATTCATTCAGATGC 4560
4501 TCTCTGTTTAAAGTACAGACACTGAGTCTTACGGAATTAATGAATTCATTCATTCAGATGC 4560
4561 CTTTATGTTGCTGTATTGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
4561 CTTTATGTTGCTGTATTGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
4621 TGTAGTCAATTCAGATTTTCAAGATATTCATATTCATATTCATATTCATATTCATATTCAT 4680
4621 TGTAGTCAATTCAGATTTTCAAGATATTCATATTCATATTCATATTCATATTCATATTCAT 4680
4681 AATCCTTTAGAGTGTCTCCCTTAAGTTTTCGGGATTTTCGGATTCAGTGTGAGTGTGAGT 4740
4681 AATCCTTTAGAGTGTCTCCCTTAAGTTTTCGGGATTTTCGGATTCAGTGTGAGTGTGAGT 4740
4741 TCTCAAGAGGCTCAGTACAGCAGCAACAAAGCAGCAAGATTTGTTACTAGTGTGTTAAACA 4800
4741 TCTCAAGAGGCTCAGTACAGCAGCAACAAAGCAGCAAGATTTGTTACTAGTGTGTTAAACA 4800
4801 ATTCCAGATGATACCAATGTCCTGCTATCTCTTCCGAAAGAGTCTTCAAGATTTAGTTC 4860
4801 ATTCCAGATGATACCAATGTCCTGCTATCTCTTCCGAAAGAGTCTTCAAGATTTAGTTC 4860

QY 4861 ATTAATTTGAGTGTCTCACTTCACTAAATCTCATGAGCTGGGCTTTTAAACAATTAAG 4920
Db 4861 ATTAATTTGAGTGTCTCACTTCACTAAATCTCATGAGCTGGGCTTTTAAACAATTAAG 4920
QY 4921 GAGAAGTATTCCTCAAAACATTTGATGATGATGATGATGATGATGATGATGATGATGATG 4980
Db 4921 GAGAAGTATTCCTCAAAACATTTGATGATGATGATGATGATGATGATGATGATGATGATG 4980
QY 4981 TCACACTGACATTCAGACTTCCGCTGCGAGGTTTCATCAAGAAATTTTCAAGATGTA 5040
Db 4981 TCACACTGACATTCAGACTTCCGCTGCGAGGTTTCATCAAGAAATTTTCAAGATGTA 5040
QY 5041 CAGTTTCTCAAAATGATGAAAGCAGAGGCTGTGTGGCAACACCAACCAAGCAACCT 5100
Db 5041 CAGTTTCTCAAAATGATGAAAGCAGAGGCTGTGTGGCAACACCAACCAAGCAACCT 5100
QY 5101 ATAGTTGATACATCTGCTGAATCTGTA 5127
Db 5101 ATAGTTGATACATCTGCTGAATCTGTA 5127
RESULT 2
ID ADR44434 standard; cDNA; 5662 BP.
XX ADR44434;
AC ADR44434;
DT 18-NOV-2004 (first entry)
DE Human p200 encoding cDNA SEQ ID NO:1.
KW mTOR-associated protein; mTOR-AP; cytostatic; antidiabetic; gene therapy;
KW mTOR-AP modulation; cancer; tumour; diabetes; human; p200; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 28.5154
CDS /*tag= a
FT /product= "p200"
XX WO2004074448-A2.
XX 02-SEP-2004.
XX 18-FEB-2004; 2004WO-US004821.
XX 18-FEB-2003; 2003US-048035P.
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX Sabatini DM, Kim D, Sarbassov DD; ✓
XX WPI; 2004-635565/61.
XX P-PSDB; ADR44436.
XX New mTOR-associated protein (mTOR-AP) polypeptide, useful for treating or
XX preventing a disorder that is responsive to mTOR-AP modulation, e.g.
XX cancer or diabetes.
XX Claim 14; SEQ ID NO 1; 91pp; English.
XX The present invention describes an isolated mTOR-associated protein (mTOR
XX -AP) polypeptide comprising a sequence that is 70% identical to the 1708
XX amino acid sequence of SEQ ID NO:3 (ADR44436, P1) or 70% identical to the
XX 327 amino acid sequence of SEQ ID NO:6 (ADR44439, P2), or that is encoded
XX by a nucleic acid that hybridises under high stringency conditions to a
XX complement sequence of the 951271 nucleic acid sequence of SEQ ID NO:2
XX (ADR44435, S1) or the 981 nucleic acid sequence of SEQ ID NO:5 (ADR44438,
XX S2). Also described: (1) an isolated antibody, or its fragment, which is
XX specifically immunoreactive with an epitope of an amino acid sequence of
XX P1 or P2; (2) the isolated nucleic acid encoding the polypeptide, or its

complement; (3) a vector comprising a nucleic acid sequence encoding an mTOR-AP polypeptide; (4) an isolated host cell comprising the nucleic acid; (5) a method of producing an mTOR-AP polypeptide; (6) a method for detecting the presence of an mTOR-AP polypeptide; (7) a kit for detecting an mTOR-AP polypeptide comprising the antibody and a detectable label for detecting the antibody; (8) a method for detecting the presence of the nucleic acid in a sample; (9) a kit comprising the probe or primer nucleic acid and instructions for use; (10) an isolated, purified or recombinant complex comprising an mTOR polypeptide and an mTOR-AP; (11) a method for identifying a compound which modulates activity of an mTOR-AP polypeptide; (12) a method of inhibiting aberrant activity of an mTOR-AP-expressing cell; (13) a method of treating or preventing a disorder that is responsive to mTOR-AP modulation, in a subject; and (14) a transgenic mouse having germline and somatic cells comprising a chromosomally incorporated transgene that disrupts the genomic mTOR-AP gene and inhibits expression of the gene, where the disruption comprises insertion of a selectable marker sequence resulting in the transgenic mouse exhibiting increased susceptibility to the formation of tumours as compared to the wild type mouse. mTOR-AP sequences cytotatic and antidiabetic activities, and can be used in gene therapy. The polypeptides, polynucleotides, compounds and methods are useful for treating or preventing a disorder that is responsive to mTOR-AP modulation, e.g. cancer or diabetes. The compound, which modulates an mTOR-AP activity or expression is useful for the manufacture of a medicament for treating a disease affected by aberrant mTOR-AP activity or expression. The present sequence encodes human p200, which is an mTOR-AP used in the exemplification of the present invention.

XX SQ Sequence 5662 BP; 1786 A; 1128 C; 1156 G; 1592 T; 0 U; 0 Other;

Query Match 100.0%; Score 5127; DB 13; Length 5662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCGGATCGCGCGCGCGCTCTGTAAGAACCTCCGAGTACGAGGCGCGAATGAC 60
Db 28 ATGCGCGGATCGCGCGCGCGCTCTGTAAGAACCTCCGAGTACGAGGCGCGAATGAC 87

Qy 61 AGCGCGGAGGAGACGTCGCCGCTGGATCTGACCCGAGAACCTCTGATAACTTAAGAGAG 120
Db 88 AGCGCGGAGGAGACGTCGCCGCTGGATCTGACCCGAGAACCTCTGATAACTTAAGAGAG 147

Qy 121 ATTCTCCAAATGTGGCGAGATTCGAGGAGTATCAATATGAGAAAGCTAGGCCATCTG 180
Db 148 ATTCTCCAAATGTGGCGAGATTCGAGGAGTATCAATATGAGAAAGCTAGGCCATCTG 207

Qy 181 AATAACTTACTAGACTCTTTGTGATATTGGCCACAGTGAAGAAACTGGGCTTTTAC 240
Db 208 AATAACTTACTAGACTCTTTGTGATATTGGCCACAGTGAAGAAACTGGGCTTTTAC 267

Qy 241 TATGAGGATATCAATAATTTGTTGCGGTAGCTTTTAAATGAAGCAAGAAAGTGCGA 300
Db 268 TATGAGGATATCAATAATTTGTTGCGGTAGCTTTTAAATGAAGCAAGAAAGTGCGA 327

Qy 301 GCAGCAGGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGATTTCTCCAGAGGTG 360
Db 328 GCAGCAGGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGATTTCTCCAGAGGTG 387

Qy 361 CTAAATTTGAAGTGGACTATTAAATAGCTAGGTGCAATGACATACACGAGCAACGAG 420
Db 388 CTAAATTTGAAGTGGACTATTAAATAGCTAGGTGCAATGACATACACGAGCAACGAG 447

Qy 421 GTAGAGAGGACACAGCACTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG 480
Db 448 GTAGAGAGGACACAGCACTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG 507

Qy 481 TTTCTTAGTTCGTGACCACTCAATTAATTTGAGTTGGAATATGATGGACTTTCAAGAAAGA 540
Db 508 TTTCTTAGTTCGTGACCACTCAATTAATTTGAGTTGGAATATGATGGACTTTCAAGAAAGA 567

Qy 541 GACAGATGTGCCAGCATGCAATTTGCAATTTATCTGTGAACCTAGACATTCAGATTCACAG 600
Db 568 GACAGATGTGCCAGCATGCAATTTGCAATTTATCTGTGAACCTAGACATTCAGATTCACAG 627

Qy 601 GTGCTGGCCCTTCGAGGAGCACTAAACACCATATTGAAAAATGTGATTGATTCACAAATTA 660
Db 628 GTGCTGGCCCTTCGAGGAGCACTAAACACCATATTGAAAAATGTGATTGATTCACAAATTA 687

Qy 661 AGTCGAATAAATGAGGCGCTTAATTAATTAATTTTGCACCTTCTTAATCATCCAAAGACT 720
Db 688 AGTCGAATAAATGAGGCGCTTAATTAATTAATTTTGCACCTTCTTAATCATCCAAAGACT 747

Qy 721 CGGCAGTATGTGCGAGCTGATGATAGAAATTAGAGAGAAATTTTGTAGCACCTTACTGATTTT 780
Db 748 CGGCAGTATGTGCGAGCTGATGATAGAAATTAGAGAGAAATTTTGTAGCACCTTACTGATTTT 807

Qy 781 CACTACAGACATAGTCCAGATACAGCTCAAGGACAGCTCAAGAGAGACAGAGAGACGGA 840
Db 808 CACTACAGACATAGTCCAGATACAGCTCAAGGACAGCTCAAGAGAGACAGAGAGACGGA 867

Qy 841 TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGATCATCGGCGAGGTATTATTAAAT 900
Db 868 TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGATCATCGGCGAGGTATTATTAAAT 927

Qy 901 TTATGTAACCTGGAATTTCTGGGATCCAGTCTCTAATAGGAGTACTTTGTCATACCAAT 960
Db 928 TTATGTAACCTGGAATTTCTGGGATCCAGTCTCTAATAGGAGTACTTTGTCATACCAAT 987

Qy 961 ATGGAATAAAGCGGAGGTCTACTTTGAAGTCTTTATGATATATTTCGTCTTCTCTACCT 1020
Db 988 ATGGAATAAAGCGGAGGTCTACTTTGAAGTCTTTATGATATATTTCGTCTTCTCTACCT 1047

Qy 1021 GTTGTGATCTGAGGAGTTCTAGAGACACTACTAGTGTAGATCCAGGAGGTTCCAAAGAC 1080
Db 1048 GTTGTGATCTGAGGAGTTCTAGAGACACTACTAGTGTAGATCCAGGAGGTTCCAAAGAC 1107

Qy 1081 AGTTGGAGGCTTTTCAGATGGCTTTGTCGAGCTGAGGCAAAACTATTCTTCTCTCATCGT 1140
Db 1108 AGTTGGAGGCTTTTCAGATGGCTTTGTCGAGCTGAGGCAAAACTATTCTTCTCTCATCGT 1167

Qy 1141 GCCAGATCCAGGCGCAGACCTCATGGATAATTTATTTGGCACTGATACCTCTCTGCAATTTAT 1200
Db 1168 GCCAGATCCAGGCGCAGACCTCATGGATAATTTATTTGGCACTGATACCTCTCTGCAATTTAT 1227

Qy 1201 CGTAATGGACTTTTGTAGAGGCTTAGTTGAAAGTGAATAACAAACAGTATGATCATCTCA 1260
Db 1228 CGTAATGGACTTTTGTAGAGGCTTAGTTGAAAGTGAATAACAAACAGTATGATCATCTCA 1287

Qy 1261 GTTAGAGCTACCACTCTTTTGTAGGAGCTTTTACATATGCAACACAAATTTCTTCTCCAT 1320
Db 1288 GTTAGAGCTACCACTCTTTTGTAGGAGCTTTTACATATGCAACACAAATTTCTTCTCCAT 1347

Qy 1321 TCACATAGCCATCAATTTTACACTGCTTCCCAACCTAATGATGATGGCTGCAATCCTTTGAT 1380
Db 1348 TCACATAGCCATCAATTTTACACTGCTTCCCAACCTAATGATGATGGCTGCAATCCTTTGAT 1407

Qy 1381 ATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCTTGAACCTGTTTTAAACGCTTCCAT 1440
Db 1408 ATCCCAAGGAAAGAGACTGCGAGCCAGTGCAGCTTGAACCTGTTTTAAACGCTTCCAT 1467

Qy 1441 GAAATGAGAAACGAGGACCTTAAGCTTTATAGTCTTCAATTTAGACCAATTTTTCAGAAA 1500
Db 1468 GAAATGAGAAACGAGGACCTTAAGCTTTATAGTCTTCAATTTAGACCAATTTTTCAGAAA 1527

Qy 1501 GCAATTCACACACACGAGGATCAGTATCTCCGAGTTTCAGAAAGATATATTTATC 1560
Db 1528 GCAATTCACACACACGAGGATCAGTATCTCCGAGTTTCAGAAAGATATATTTATC 1587

Qy 1561 CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGCTCTTCAACAT 1620
Db 1588 CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGCTCTTCAACAT 1647

Qy 1621 AAAGAGATCTTGAATGGAATTTGAATCTTATAGGAGCAATTTCTTAAGTGCCCAAAATGTA 1680
Db 1648 AAAGAGATCTTGAATGGAATTTGAATCTTATAGGAGCAATTTCTTAAGTGCCCAAAATGTA 1707

Db 3868 AATCAAAATTCGGTGTCCCTGGTCCAGGTTCTTCTCATACGCTTCTTAGAAGAGCA 3927
Qy 3901 CAGTCCCTTAAGACACCCCTCTATTGTCTACAAATTAAGAGTCTAGCAGATGTGAACTTTAGT 3960
Db 3928 CAGTCCCTTAAGACACCCCTCTATTGTCTACAAATTAAGAGTCTAGCAGATGTGAACTTTAGT 3987
Qy 3961 TACAAGAGTCTAGAGATGCTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAGA 4020
Db 3988 TACAAGAGTCTAGAGATGCTTTTGGCTATGCTACACTGAAAGACTACAGCAACAAGA 4047
Qy 4021 ATGCATCCATCCTTATCTCACTCTGAAGCTTTGCGATCTCCAGCAAAAAGATGCTCTATTT 4080
Db 4048 ATGCATCCATCCTTATCTCACTCTGAAGCTTTGCGATCTCCAGCAAAAAGATGCTCTATTT 4107
Qy 4081 ACTGATACCATCAACATCAAGGCCAACAGATTTTTCAGTCCAGATTAACACCAAGCAGGTTTC 4140
Db 4108 ACTGATACCATCAACATCAAGGCCAACAGATTTTTCAGTCCAGATTAACACCAAGCAGGTTTC 4167
Qy 4141 ATGAAGCCTTAAGTTATGATCATATAGATAAAGAGATTTATTTAGTCTCTATTAATCAA 4200
Db 4168 ATGAAGCCTTAAGTTATGATCATATAGATAAAGAGATTTATTTAGTCTCTATTAATCAA 4227
Qy 4201 AATACCCCTGCAACCATCTTCTCAGTGGGTCATGTGTCAGTGCACATATGCGGGT 4260
Db 4228 AATACCCCTGCAACCATCTTCTCAGTGGGTCATGTGTCAGTGCACATATGCGGGT 4287
Qy 4261 TCAGATGATTAACATTTGGTCTTCTCCTCCGCTGATATATAATGATATATTTCCAGGTAAG 4320
Db 4288 TCAGATGATTAACATTTGGTCTTCTCCTCCGCTGATATATAATGATATATTTCCAGGTAAG 4347
Qy 4321 GATATCCCTTATTTTCAGACAAAACATACACCAATGATGATCGAGGTGCAAGAGCA 4380
Db 4348 GATATCCCTTATTTTCAGACAAAACATACACCAATGATGATCGAGGTGCAAGAGCA 4407
Qy 4381 TTTGCCATGATGACAGAGGCTTCTCCATCTGGAATCGAGGCTCTTGAATAAATCTCTTT 4440
Db 4408 TTTGCCATGATGACAGAGGCTTCTCCATCTGGAATCGAGGCTCTTGAATAAATCTCTTT 4467
Qy 4441 CACTTGCTACGACAGCAGATGAGTCTTACGGAATAATGAATCAATCCATTCAGATGCC 4500
Db 4468 CACTTGCTACGACAGCAGATGAGTCTTACGGAATAATGAATCAATCCATTCAGATGCC 4527
Qy 4501 TCTCTGTTTTAGAAAGTACAGAAAGACATGGAATCAGGAACATACAGATGATACTGC 4560
Db 4528 TCTCTGTTTTAGAAAGTACAGAAAGACATGGAATCAGGAACATACAGATGATACTGC 4587
Qy 4561 CTTTATTTGCTGCTATTGAAATCTGGGTTTCCAGCCAGCACCACTGAGTGCATA 4620
Db 4588 CTTTATTTGCTGCTATTGAAATCTGGGTTTCCAGCCAGCACCACTGAGTGCATA 4647
Qy 4621 TGTAGTCATTCAGACTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4680
Db 4648 TGTAGTCATTCAGACTTTCAAGATATTCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4707
Qy 4681 AATCTTTTAAAGTGGTCCCTCTAAGTTTTCGGGATTTCTGATCAGTGTGAGGTTG 4740
Db 4708 AATCTTTTAAAGTGGTGGCTCTAAGTTTTCGGGATTTCTGATCAGTGTGAGGTTG 4767
Qy 4741 TCTCAAGAGGCTCAGTACAGCAGCAACCAAGCAGCAATTTTACTAGTGTAAACA 4800
Db 4768 TCTCAAGAGGCTCAGTACAGCAGCAACCAAGCAGCAATTTTACTAGTGTAAACA 4827
Qy 4801 ATTCCAGATGATACCAATGTCGGCTATCTCTCTCGCAAGAGGTTCTAAGATTAGTC 4860
Db 4828 ATTCCAGATGATACCAATGTCGGCTATCTCTCTCGCAAGAGGTTCTAAGATTAGTC 4887
Qy 4861 ATTAATTTAGTAGTTCAGTTTCAACTTAATGTCATGAGACTGGGCTTTTAAACAATTAAG 4920
Db 4888 ATTAATTTAGTAGTTCAGTTTCAACTTAATGTCATGAGACTGGGCTTTTAAACAATTAAG 4947
Qy 4921 GAGAGTATCTCTAAACATTTGATGACATATGCTTTTACTCTGAGGTTTCCCATTTGCTG 4980
Db 4948 GAGAGTATCTCTAAACATTTGATGACATATGCTTTTACTCTGAGGTTTCCCATTTGCTG 5007

Qy 4981 TCACACTGCACATTCAGACTTCCGTGTCGAGGTTTCATACAGAAATTTATTTCAAGATGTA 5040
Db 5008 TCACACTGCACATTCAGACTTCCGTGTCGAGGTTTCATACAGAAATTTATTTCAAGATGTA 5067
Qy 5041 CAGTTTCTACAAATGCAATGAAGAGCAGAGCTGTGTGGCAACACCAAGCAACCT 5100
Db 5068 CAGTTTCTACAAATGCAATGAAGAGCAGAGCTGTGTGGCAACACCAAGCAACCT 5127
Qy 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127
Db 5128 ATAGTTGATACATCTGCTGAATCCTGA 5154

RESULT 3

ID ADQ89851 standard; DNA; 3829 BP.

XX ADQ89851;

AC AC

XX 21-OCT-2004 (first entry) ✓

XX Antagonist of cell cycle progression nucleotide sequence #141.

XX Cytostatic; cancer; cell division cycle; mitosis; meiosis;

XX cell cycle progression; ds.

XX Homo sapiens.

XX WO2004063362-A2.

XX 29-JUL-2004.

XX 31-DEC-2003. 2003WO-GB005635.

XX 10-JAN-2003; 2003US-0439123P.

XX 06-MAY-2003; 2003US-0468402P.

XX (CVCL-) CYCLACEL LTD.

XX Glover D, Bell G, Frenz L, Midgley C;

XX WPI; 2004-544089/52.

XX P-PSDB; ADQ89852.

XX New cell cycle progression genes and proteins for modulating cell cycle

XX progression in cells, for preventing, treating or diagnosing cell

XX proliferative diseases (e.g. cancer) or for identifying modulators of

XX mitosis or meiosis.

XX Claim 1; SEQ ID NO 281; 461pp; English.

XX The present invention relates to a polynucleotide for preventing,

XX treating or diagnosing a disease in an individual. The composition or the

XX polypeptide, polynucleotide or RNA precursor, or antibody is useful for

XX diagnosing, preventing or treating diseases (e.g. cell proliferative

XX diseases such as cancer) in an individual. These may also be used for

XX identifying substances capable of binding to or modulating the function

XX of the polypeptide, capable of affecting the function of the

XX corresponding gene, or capable of inhibiting the cell division cycle or

XX cell cycle progression, preferably mitosis and/or meiosis. The present

XX sequence represents an antagonist of cell cycle progression nucleotide

XX sequence.

XX SQ Sequence 3829 BP; 1230 A; 776 C; 748 G; 1075 T; 0 U; 0 Other;

Query Match

Best Local Similarity 74.7%; Score 3829; DB 13; Length 3829;

Matches 3829; Conservativity 100.0%; Pred. No. 0;

Mismatches 0; Indels 0; Gaps 0;

Qy 1299 GGCAAAACAAATTTCTTCTCTCATTCACATAGCCATCATTTACATGCTTGCACACCTTAAT 1358
Db 1 GGCAAAACAAATTTCTTCTCTCATTCACATAGCCATCATTTACATGCTTGCACACCTTAAT 60

QY 1359 GAATATGGCTGCATCCTTTTGATATCCCAAGAAAGAGACTGCGAGCCAGTGCAGCCTT 1418
DB 61 GAAATATGGCTGCATCCTTTTGATATCCCAAGAAAGAGACTGCGAGCCAGTGCAGCCTT 120
QY 1419 GAACTGTTTAAACCGCTTCCATGAATGAAGAACGAGGACCTTAAGCCCTTATAGTCTTCA 1478
DB 121 GAACTGTTTAAACCGCTTCCATGAATGAAGAACGAGGACCTTAAGCCCTTATAGTCTTCA 180
QY 1479 TTTAGACCACTATTTTCAGAAAGCAATTCGAACACACAGGAAACCGGATCAGTATCTCCG 1538
DB 181 TTTAGACCACTATTTTCAGAAAGCAATTCGAACACACAGGAAACCGGATCAGTATCTCCG 240
QY 1539 AGTTTCAGAAAGATATATTTTATCTTTAAGGATACAGAGGAGCTCTTTTAAATAACCTTAG 1598
DB 241 AGTTTCAGAAAGATATATTTTATCTTTAAGGATACAGAGGAGCTCTTTTAAATAACCTTAG 300
QY 1599 AGATAGCCAAAGTCTTCAACATTAAGAGNATCTTGAATGGAAATTCGAATCTTATAGGCAC 1658
DB 301 AGATAGCCAAAGTCTTCAACATTAAGAGNATCTTGAATGGAAATTCGAATCTTATAGGCAC 360
QY 1659 CATTTCTTAAGTGGCCAAATGTAATCTAAGAACTATTAAGATGAACAGTTCACAGGTT 1718
DB 361 CATTTCTTAAGTGGCCAAATGTAATCTAAGAACTATTAAGATGAACAGTTCACAGGTT 420
QY 1719 TGTACGAAGACTACTTTTATTTTACAGCCAGCAGTAAATTAATATGCAACCTGGATCT 1778
DB 421 TGTACGAAGACTACTTTTATTTTACAGCCAGCAGTAAATTAATATGCAACCTGGATCT 480
QY 1779 GGAATTTGCAAGGCCAAGCCTACAGCTGAGTGGCCAGTTCACAGAAATTTCTTCT 1838
DB 481 GGAATTTGCAAGGCCAAGCCTACAGCTGAGTGGCCAGTTCACAGAAATTTCTTCT 540
QY 1839 TGAATCTGAAGAGGATGGGCAAGGCTACTTAGAAGATCTAGTAAAGGATATTTGTTCAAGTG 1898
DB 541 TGAATCTGAAGAGGATGGGCAAGGCTACTTAGAAGATCTAGTAAAGGATATTTGTTCAAGTG 600
QY 1899 GCTCAATGCTTCACTGGAATGAACCCGGAAGAGCTTCAAAATATGTTTATTTGAC 1958
DB 601 GCTCAATGCTTCACTGGAATGAACCCGGAAGAGCTTCAAAATATGTTTATTTGAC 660
QY 1959 CACCCCTTAGTCAACACTACTTTTATTTATTTGAAACACTTTCTTGCCACCCCTCATGGAGT 2018
DB 661 CACCCCTTAGTCAACACTACTTTTATTTATTTGAAACACTTTCTTGCCACCCCTCATGGAGT 720
QY 2019 TAAAAATGCTGGAAAAATGCAAGTGAATTCAGTGTCTCTTAATCTTTGCTTCTTGAANA 2078
DB 721 TAAAAATGCTGGAAAAATGCAAGTGAATTCAGTGTCTCTTAATCTTTGCTTCTTGAANA 780
QY 2079 CCAAGATCACTTGTAAACCTTACTGTTTCTAGCTTGGACTATAGCAGAGATGGATTGGC 2138
DB 781 CCAAGATCACTTGTAAACCTTACTGTTTCTAGCTTGGACTATAGCAGAGATGGATTGGC 840
QY 2139 TAGAGTCACTCTTTCCAAAAATTTTAACTGCAGCTACTGATGCTGCAAGCTCTATGCAAC 2198
DB 841 TAGAGTCACTCTTTCCAAAAATTTTAACTGCAGCTACTGATGCTGCAAGCTCTATGCAAC 900
QY 2199 AAAACATTTAAGGGTATATTGAGAGCTAATGTTGAATCTTTTAAATTTGGGGAAATTGA 2258
DB 901 AAAACATTTAAGGGTATATTGAGAGCTAATGTTGAATCTTTTAAATTTGGGGAAATTGA 960
QY 2259 GTTGTAGTGAACCCAGCTACATGATAAAAAACAAACGATTTCTCTGAAGCTCTTGATAT 2318
DB 961 GTTGTAGTGAACCCAGCTACATGATAAAAAACAAACGATTTCTCTGAAGCTCTTGATAT 1020
QY 2319 CCTCGATGAAGCATGTGAAGCAAGGCCAATCTTTCTGCTCTCATTTGAGATGAACCCAGC 2378
DB 1021 CCTCGATGAAGCATGTGAAGCAAGGCCAATCTTTCTGCTCTCATTTGAGATGAACCCAGC 1080
QY 2379 GTTATCCACCTTGGAGCAAGGGTTGCTTCTCTGCTGAGATTTCTCTCCATTCCAA 2438
DB 1081 GTTATCCACCTTGGAGCAAGGGTTGCTTCTCTGCTGAGATTTCTCTCCATTCCAA 1140

QY 2439 AGGATTTTCTCTATCTGAATGAAAGAGGTTATGTAGCAAAACAAATTTGAAAGTGGACAG 2498
DB 1141 AGGATTTTCTCTATCTGAATGAAAGAGGTTATGTAGCAAAACAAATTTGAAAGTGGACAG 1200
QY 2499 GGAATACAACTCCAAATATGTTGACTTGAATGAGGAAACAACTCAATGAAGCACTTACTAC 2558
DB 1201 GGAATACAACTCCAAATATGTTGACTTGAATGAGGAAACAACTCAATGAAGCACTTACTAC 1260
QY 2559 TTACCGGAAGCCTGTTGATGCTGATTAATGTTTCTGCGAGTAAACCAAGATTTACAGCG 2618
DB 1261 TTACCGGAAGCCTGTTGATGCTGATTAATGTTTCTGCGAGTAAACCAAGATTTACAGCG 1320
QY 2619 TCCTCAGCTCTACTGCTATACACCTTTATGACAACTAGTACACCATAAAAACAGCCTG 2678
DB 1321 TCCTCAGCTCTACTGCTATACACCTTTATGACAACTAGTACACCATAAAAACAGCCTG 1380
QY 2679 CCATTTGTTGGAAGTACAGAAATATTTACAGAACTCTGTCGTAAATGTTTCTGTACACAGA 2738
DB 1381 CCATTTGTTGGAAGTACAGAAATATTTACAGAACTCTGTCGTAAATGTTTCTGTACACAGA 1440
QY 2739 TTTTGGATAAGTGGGAAGAAATTTAAAAAATCTGAAAGCATCTCTTTGGGCCCTTGGGAATAT 2798
DB 1441 TTTTGGATAAGTGGGAAGAAATTTAAAAAATCTGAAAGCATCTCTTTGGGCCCTTGGGAATAT 1500
QY 2799 CGGCTCATCAAAATTTGGGCTCTCAATTTGCTACAGAAAGAAACGTTGATTCAGATATACT 2858
DB 1501 CGGCTCATCAAAATTTGGGCTCTCAATTTGCTACAGAAAGAAACGTTGATTCAGATATACT 1560
QY 2859 AAAAATTTGCAAAAAACAGTGTGAAGTCTTTCCATCAGAGGACCTGTGTATATGACTTGG 2918
DB 1561 AAAAATTTGCAAAAAACAGTGTGAAGTCTTTCCATCAGAGGACCTGTGTATATGACTTGG 1620
QY 2919 GCTCATAGTAAAAACCAACCAAGCTGTGATATTTCTAAAAATGTCAAACTGGGATGCTGT 2978
DB 1621 GCTCATAGTAAAAACCAACCAAGCTGTGATATTTCTAAAAATGTCAAACTGGGATGCTGT 1680
QY 2979 GAGGCATAGTCCGAAACATCTGTGGCCAGTGGTTCAGATGATGTGGAACAACTCTGTAA 3038
DB 1681 GAGGCATAGTCCGAAACATCTGTGGCCAGTGGTTCAGATGATGTGGAACAACTCTGTAA 1740
QY 3039 TGAATTTTCACTATCCCAAGCCTCTAAGTTTGAATTCGAGTCAACAGCTCTAGACA 3098
DB 1741 TGAATTTTCACTATCCCAAGCCTCTAAGTTTGAATTCGAGTCAACAGCTCTAGACA 1800
QY 3099 TAAATAGTGAAAGTGAATCTGTGCCATCGAGTATGTTCAATATGGAAGATGACCGGTTGG 3158
DB 1801 TAAATAGTGAAAGTGAATCTGTGCCATCGAGTATGTTCAATATGGAAGATGACCGGTTGG 1860
QY 3159 CAGCAGCTCTACTAGTACATTTTCTGATATCAATGAAGATACAGAGCCAAACATTTTA 3218
DB 1861 CAGCAGCTCTACTAGTACATTTTCTGATATCAATGAAGATACAGAGCCAAACATTTTA 1920
QY 3219 TCACCCGATCTGACCCATTAAGGATTAATAATTTCAATTCCTTTTCTTTGCTCTTAGTAAACT 3278
DB 1921 TCACCCGATCTGACCCATTAAGGATTAATAATTTCAATTCCTTTTCTTTGCTCTTAGTAAACT 1980
QY 3279 TGTGAAGATCGTATCTTAAATTCGCTTAACTTTGCTTAAACAAACATCTGCTAGTAGCAG 3338
DB 1981 TGTGAAGATCGTATCTTAAATTCGCTTAACTTTGCTTAAACAAACATCTGCTAGTAGCAG 2040
QY 3339 TGAATCCAAAGAGGGAAATTTATCATCTGAAAGTAAAGCAACAGCCGAATCAGAAC 3398
DB 2041 TGAATCCAAAGAGGGAAATTTATCATCTGAAAGTAAAGCAACAGCCGAATCAGAAC 2100
QY 3399 ACTTACCGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACCATATCCACTGT 3458
DB 2101 ACTTACCGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACCATATCCACTGT 2160
QY 3459 ACAGAAAAATTTACAAATTAAGACTTTTATGGGATTAAGCACTTTGAAGACACTGG 3518
DB 2161 ACAGAAAAATTTACAAATTAAGACTTTTATGGGAAATTAAGCACTTTGAAGACACTGG 2220
QY 3519 TAGTACCAACAGCATTTGGAGAAAAATGACTTAAAAATTCACCAAGAAATTTTGGTACAGAGAA 3578

Db 2221 TAGTACACCAAGCATTTGGAGAAATGACTTAAATTCACCAAGAAATTTGGTCAGAGAA 2280
Qy 3579 TCACAGAGAAATACAAAGCCGAGAGAGTGTAGTGTAGAAAGTTCAACGAGCTCACATAT 3638
Db 2281 TCACAGAGAAATACAAAGCCGAGAGAGTGTAGTGTAGAAAGTTCAACGAGCTCACATAT 2340
Qy 3639 GAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATTAAGTTCAATGAG 3698
Db 2341 GAAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATTAAGTTCAATGAG 2400
Qy 3699 CTCAGTCCCTTCACGAGAGACAGTAGGTGTAGTGTACAACTATGACACAGACTGTGG 3758
Db 2401 CTCAGTCCCTTCACGAGAGACAGTAGGTGTAGTGTACAACTATGACACAGACTGTGG 2460
Qy 3759 AAGCATGAGTACTGTGGTAAAGTACTAAAGTCTAAAGTCTAAAGTCTAAAGTCTAAAGTCTAAAG 3818
Db 2461 AAGCATGAGTACTGTGGTAAAGTACTAAAGTCTAAAGTCTAAAGTCTAAAGTCTAAAGTCTAAAG 2520
Qy 3819 GTCTAACCATCTGTCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGTCCCTGGTGTCTTC 3878
Db 2521 GTCTAACCATCTGTCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGTCCCTGGTGTCTTC 2580
Qy 3879 TCATACGCTTCTTAGAAGACAGTCCCTTAAAGCACCTCTATTTGCTACAAATTAAG 3938
Db 2581 TCATACGCTTCTTAGAAGACAGTCCCTTAAAGCACCTCTATTTGCTACAAATTAAG 2640
Qy 3939 TCTAGCAGATTGTAACTTTTACACAAAGTCTTACAGAGTGTCTTGGCTATGCTACACT 3998
Db 2641 TCTAGCAGATTGTAACTTTTACACAAAGTCTTACAGAGTGTCTTGGCTATGCTACACT 2700
Qy 3999 GAAAAGACTACAGCAACAAAGAAATGCATCCATCTCTCACTCTGAAGCTTTGGCATC 4058
Db 2701 GAAAAGACTACAGCAACAAAGAAATGCATCCATCTCTCACTCTGAAGCTTTGGCATC 2760
Qy 4059 TCACAGAAAGATGTGCTATTTACTGATACATCACATGAAAGCCCAACAGTTTGGTGC 4118
Db 2761 TCACAGAAAGATGTGCTATTTACTGATACATCACATGAAAGCCCAACAGTTTGGTGC 2820
Qy 4119 CAGATTAAACCAAGCAGGTTTCATGAAGCTTAAAGTTATGTCATCATTAAGTAAAGAA 4178
Db 2821 CAGATTAAACCAAGCAGGTTTCATGAAGCTTAAAGTTATGTCATCATTAAGTAAAGAA 2880
Qy 4179 TTTATTGAGTCTTAAATCAAAATACCTGCAACGATCTTCTCAGTGGCGTCCATGGT 4238
Db 2881 TTTATTGAGTCTTAAATCAAAATACCTGCAACGATCTTCTCAGTGGCGTCCATGGT 2940
Qy 4239 GTCCAGTGCCACATATGGGGGTTCAGATGATTAATGGTCTTGGCTCTCCCGGTGGATAT 4298
Db 2941 GTCCAGTGCCACATATGGGGGTTCAGATGATTAATGGTCTTGGCTCTCCCGGTGGATAT 3000
Qy 4299 AAATGATATATCCAGTAAAGATATTCCTTATTCAGCAAAACAAACATACCAACA 4358
Db 3001 AAATGATATATCCAGTAAAGATATTCCTTATTCAGCAAAACAAACATACCAACA 3060
Qy 4359 TGATGATCGAGGTGCAAGGATATTCCTTATTCAGCAAAACAAACATACCAACA 4418
Db 3061 TGATGATCGAGGTGCAAGGATATTCCTTATTCAGCAAAACAAACATACCAACA 3120
Qy 4419 AGTCTTTGTAATAAATTTCTTTTCACTTGTGCTACAGCAGATGAGTCTTACGGAATAAT 4478
Db 3121 AGTCTTTGTAATAAATTTCTTTTCACTTGTGCTACAGCAGATGAGTCTTACGGAATAAT 3180
Qy 4479 GAATTCATCCATTCAGTCCCTCTGTTTTTGTAGAAAGTACAGAGACACTGACTACA 4538
Db 3181 GAATTCATCCATTCAGTCCCTCTGTTTTTGTAGAAAGTACAGAGACACTGACTACA 3240
Qy 4539 GGAACATACAGATGAATACCTCTTATTTGCTGTGCTGATTAAGAAATCTGGGTTTCCAGCC 4598
Db 3241 GGAACATACAGATGAATACCTCTTATTTGCTGTGCTGATTAAGAAATCTGGGTTTCCAGCC 3300
Qy 4599 CAGCAACCAACTGAGTGCATATGTAGTCAATTCAGACTTTTCAAGATATTCATATTTCTGA 4658

Db 3301 CAGCAACCAACTGAGTGCATATGTAGTCAATTCAGACTTTCAAGATATTCATATTTCTGA 3360
Qy 4659 TTGTGTGAGCAGACTATCCATTAATCTTTAGAGTGTCTCCCTCTAAGTTTTCGGGGAT 4718
Db 3361 TTGTGTGAGCAGACTATCCATTAATCTTTAGAGTGTCTCCCTCTAAGTTTTCGGGGAT 3420
Qy 4719 TTCTGGATGCACTGATGGGGTGTCTCAAGAAAGGCTCAGCTAGCAGACCCAAAGCACAGA 4778
Db 3421 TTCTGGATGCACTGATGGGGTGTCTCAAGAAAGGCTCAGCTAGCAGACCCAAAGCACAGA 3480
Qy 4779 ATTGTACTAGGTGTTAAACCAATTCAGATGATACCAATGTGCGGTATATCTCTTCG 4838
Db 3481 ATTGTACTAGGTGTTAAACCAATTCAGATGATACCAATGTGCGGTATATCTCTTCG 3540
Qy 4839 CAAAGAGTTTCTAAGATTAGTCAATTAATTTGAGTGTAGTTTCAACTAAATGTCATGA 4898
Db 3541 CAAAGAGTTTCTAAGATTAGTCAATTAATTTGAGTGTAGTTTCAACTAAATGTCATGA 3600
Qy 4899 GACTGGGCTTTTAAACCAATTAAGGAGAGTATCTCTCAAAACATTTGATGACATATGCTTTA 4958
Db 3601 GACTGGGCTTTTAAACCAATTAAGGAGAGTATCTCTCAAAACATTTGATGACATATGCTTTA 3660
Qy 4959 CTCTGAGTTTCCCATTTGCTGTCACTGACATTCAGACTTCCGTGTGCGAGGTTTCAAT 5018
Db 3661 CTCTGAGTTTCCCATTTGCTGTCACTGACATTCAGACTTCCGTGTGCGAGGTTTCAAT 3720
Qy 5019 ACAAGAAATTTTCAAGATGTACAGTTTCTTACAAATGCATGAAAGCAGAGGCTGTGT 5078
Db 3721 ACAAGAAATTTTCAAGATGTACAGTTTCTTACAAATGCATGAAAGCAGAGGCTGTGT 3780
Qy 5079 GGCAACACCAACCAAGCAACTATAGTTGATATCATCTGCTGAATCTCTGA 5127
Db 3781 GGCAACACCAACCAAGCAACTATAGTTGATATCATCTGCTGAATCTCTGA 3829

RESULT 4

AD71191
ID ADE71191 standard; DNA; 8232 BP.
XX
AC ADE71191;
XX
DT 29-JAN-2004 (first entry) ✓
XX
DE Novel human protein coding sequence #7.
XX
KW human; novel protein; drug; gene; ds.
XX
OS Homo sapiens.
XX
PN JP2002345493-A.
XX
PD 03-DEC-2002.
XX
PF 29-MAR-2001; 2002JP-00049046.
XX
PR 29-MAR-2001; 2001JP-00095524.
XX
PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.
XX
DR WPI; 2003-460885/44.
DR P-PSDB; ADE71253.
XX
PT A gene and a protein encoded by it, used in drugs.
XX
PS Claim 1; SEQ ID NO 7; 257pp; Japanese.
CC The invention comprises the amino acid and coding sequences of novel
CC human proteins. The DNA and protein sequences of the invention are used
CC in drugs. The present DNA sequence encodes a novel human protein of the
CC invention.
XX
SQ Sequence 8232 BP; 2655 A; 1442 C; 1476 G; 2659 T; 0 U; 0 Other;

Query Match 74.1%; Score 3800; DB 10; Length 8232; Best Local Similarity 99.5%; Pred. No. 0; Matches 3829; Conservative 0; Mismatches 0; Indels 19; Gaps 1;									
QY	1299	GGCAACCAATCTCTCCTCATTACATAGCCATCATTTACATGCTTGGCCAACTTAAT	1358						
DB	1	GGCAACCAATCTCTCCTCATTACATAGCCATCATTTACATGCTTGGCCAACTTAAT	60						
QY	1359	GAATATGGCTGCATCCTTTGATATCCCAAGGAAAGAGACTGCGAGCGAGTGCAGCCTT	1418						
DB	61	GAATATGGCTGCATCCTTTGATATCCCAAGGAAAGAGACTGCGAGCGAGTGCAGCCTT	120						
QY	1419	GAATGTTTAAACCGTTCATGAATGAAGAAACGAGGACTTAAGCTTATAGTCTTCA	1478						
DB	121	GAATGTTTAAACCGTTCATGAATGAAGAAACGAGGACTTAAGCTTATAGTCTTCA	180						
QY	1479	TTTAGACCACTATTATTCAGAAAGCAATTCGAACACACACAGAAACGGGATCAGTATCTCG	1538						
DB	181	TTTAGACCACTATTATTCAGAAAGCAATTCGAACACACACAGAAACGGGATCAGTATCTCG	240						
QY	1539	AGTTCAAGAGATATATTTATCCTTAAGGATACAGAGAGCTCTTTTAAATTAACCTTAG	1598						
DB	241	AGTTCAAGAGATATATTTATCCTTAAGGATACAGAGAGCTCTTTTAAATTAACCTTAG	300						
QY	1599	AGATAGCCAACTCCTTCAACATAGAGAACTCTTGAATGGAATGGAATCTTATAGGAC	1658						
DB	301	AGATAGCCAACTCCTTCAACATAGAGAACTCTTGAATGGAATGGAATCTTATAGGAC	360						
QY	1659	CATTCTTAAGTGGCCAAATGTAATCTAAGAACTATTAAGATGAACAGTTCACAGGTT	1718						
DB	361	CATTCTTAAGTGGCCAAATGTAATCTAAGAACTATTAAGATGAACAGTTCACAGGTT	420						
QY	1719	TGTAAGAAGACTATTTTATCAAGCCAGAGTAATTAATTAAGCAACCTGATCT	1778						
DB	421	TGTAAGAAGACTATTTTATCAAGCCAGAGTAATTAATTAAGCAACCTGATCT	480						
QY	1779	GGATTTTGCAGGCCCAACAGCTACGGTGTAGTTCGCCAGTTTACAGAAATTTCTCT	1838						
DB	481	GGATTTTGCAGGCCCAACAGCTACGGTGTAGTTCGCCAGTTTACAGAAATTTCTCT	540						
QY	1839	TGAATCTGAAGAGGATGGGCAAGGCTACTTAGAAGATCTAGTAAAGGATTTGTTCAAGT	1898						
DB	541	TGAATCTGAAGAGGATGGGCAAGGCTACTTAGAAGATCTAGTAAAGGATTTGTTCAAGT	600						
QY	1899	GCTCAATGCTTATCTGGATGAACCCGAAAGAGCTCTTCAAAATTAAGTATTTGAC	1958						
DB	601	GCTCAATGCTTATCTGGATGAACCCGAAAGAGCTCTTCAAAATTAAGTATTTGAC	660						
QY	1959	CACCTTATGTCACACTACTCTTTTATTTATTTGGAACACTTTCTTGCCACCTCATGGAGT	2018						
DB	661	CACCTTATGTCACACTACTCTTTTATTTATTTGGAACACTTTCTTGCCACCTCATGGAGT	720						
QY	2019	TAAATATGCTGGAATAATGCAAGTGTATTTCAAGTGTCTCTTAATCTTTGCTTGAATAA	2078						
DB	721	TAAATATGCTGGAATAATGCAAGTGTATTTCAAGTGTCTCTTAATCTTTGCTTGAATAA	780						
QY	2079	CCAAGATCACTTGTAAAACCTTACTCTTCTAGCTTGGACTAGCAGAGATGGATGGC	2138						
DB	781	CCAAGATCACTTGTAAAACCTTACTCTTCTAGCTTGGACTAGCAGAGATGGATGGC	840						
QY	2139	TAGAGTCACTCTTCCAAAATTTTAACTGCAGCTACTGTAGTGCCTGCAGACTCTATGCAAC	2198						
DB	841	TAGAGTCACTCTTCCAAAATTTTAACTGCAGCTACTGTAGTGCCTGCAGACTCTATGCAAC	900						
QY	2199	AAACCATTTAAGGATATTTATGAGCTAATGTTGAATCTTTTAAATTAATTTGGGGAATGA	2258						
DB	901	AAACCATTTAAGGATATTTATGAGCTAATGTTGAATCTTTTAAATTAATTTGGGGAATGA	960						
QY	2259	GTTGTTAGTCAACCCAGCTACATGATAAAACAAACGATTTCTCTGAAGCTCTTGATAT	2318						
DB	961	GTTGTTAGTCAACCCAGCTACATGATAAAACAAACGATTTCTCTGAAGCTCTTGATAT	1020						
QY	2319	CCTCGATGAAGCATGTGAAGACAAGGCCAATCTTTCATGCTCTCATTCAGATGAACAACG	2378						
DB	1021	CCTCGATGAAGCATGTGAAGACAAGGCCAATCTTTCATGCTCTCATTCAGATGAACAACG	1080						
QY	2379	GTTATCCCACTTTGGAGACAAGGTTTGTCTTCTGCTGAGATTTCTTCCATTCACAA	2438						
DB	1081	GTTATCCCACTTTGGAGACAAGGTTTGTCTTCTGCTGAGATTTCTTCCATTCACAA	1140						
QY	2439	AGGATTTTCTCTCTGAAATGAAGAGGTTATGTAGCAAAACAATTTGAAAGTGGACAG	2498						
DB	1141	AGGATTTTCTCTCTGAAATGAAGAGGTTATGTAGCAAAACAATTTGAAAGTGGACAG	1200						
QY	2499	GGATAACAATCTCCAAATATGTGATTGAGGAAACAATCTCAATGAAGCACTTACTAC	2558						
DB	1201	GGATAACAATCTCCAAATATGTGATTGAGGAAACAATCTCAATGAAGCACTTACTAC	1260						
QY	2559	TTACCGGAAGCTTGTGATGATTAATCTGTCGAGTAACCAAGATTTACAGCG	2618						
DB	1261	TTACCGGAAGCTTGTGATGATTAATCTGTCGAGTAACCAAGATTTACAGCG	1320						
QY	2619	TCCTCAGTCTTACCTGCTATACACCTTTATGGAACAACCTAGTACCAATAAAAACAGGCTG	2678						
DB	1321	TCCTCAGTCTTACCTGCTATACACCTTTATGGAACAACCTAGTACCAATAAAAACAGGCTG	1380						
QY	2679	CCATTTGTTGGAAGTACAGAATATTTACAGAACTCTGTCGTAATGTTGTTACACAGA	2738						
DB	1381	CCATTTGTTGGAAGTACAGAATATTTACAGAACTCTGTCGTAATGTTGTTACACAGA	1440						
QY	2739	TTTGGATAAGTGGGAAGAAATTAATAAATCTGAAGCACTCTTTTGGGCTTGGTTGTAA	2790						
DB	1441	TTTGGATAAGTGGGAAGAAATTAATAAATCTGAAGCACTCTTTTGGGCTTGGTTGTAA	1500						
QY	2791	-----GGAAATATCGGCTCATCAAAATCGGGTCTCAATTTGCTACAGGAAGAA	2839						
DB	1501	TAAACACTTCTGGAATATCGGCTCATCAAAATCGGGTCTCAATTTGCTACAGGAAGAA	1560						
QY	2840	ACGTGATTCAGATATATACTAAAACTTGCAAAACAGTGTGAAGTTCTTCCATCAGAGGA	2899						
DB	1561	ACGTGATTCAGATATATACTAAAACTTGCAAAACAGTGTGAAGTTCTTCCATCAGAGGA	1620						
QY	2900	CCTGTGTATATGTAATTTGGGCTCATAGTAAAAACCAACAGGCTGTGATATTTCAAAAT	2959						
DB	1621	CCTGTGTATATGTAATTTGGGCTCATAGTAAAAACCAACAGGCTGTGATATTTCAAAAT	1680						
QY	2960	GTCAAACTGGGATGCTGTGAGGATAGTCCCAAACTCTGTGGCCAGTGTTCAGATG	3019						
DB	1681	GTCAAACTGGGATGCTGTGAGGATAGTCCCAAACTCTGTGGCCAGTGTTCAGATG	1740						
QY	3020	ATGTGGAACAATCTGTAATGAATTTTCACTTATCCCAAGCACTTAAGTTTGAATCGG	3079						
DB	1741	ATGTGGAACAATCTGTAATGAATTTTCACTTATCCCAAGCACTTAAGTTTGAATCGG	1800						
QY	3080	AGTCAACAGCTCTAGACATAATAGTGAAAGTGAATCTGTGCCATCGAGTATGTTCAAT	3139						
DB	1801	AGTCAACAGCTCTAGACATAATAGTGAAAGTGAATCTGTGCCATCGAGTATGTTCAAT	1860						
QY	3140	TGGAGGATGACCGGTTTGGCAGAGCTCTAGTATCAATTTTCTTGCATCAATGAAG	3199						
DB	1861	TGGAGGATGACCGGTTTGGCAGAGCTCTAGTATCAATTTTCTTGCATCAATGAAG	1920						
QY	3200	ATACAGAGCAATTTTATGACCGATCTGACCCATTAAGGATAAAATTCATTTCCCTT	3259						
DB	1921	ATACAGAGCAATTTTATGACCGATCTGACCCATTAAGGATAAAATTCATTTCCCTT	1980						
QY	3260	TCCTTGTCTTAGTAAATCTTGTGAAGAATCGTATCTTAAATTCGTTTTCGCTAACA	3319						
DB	1981	TCCTTGTCTTAGTAAATCTTGTGAAGAATCGTATCTTAAATTCGTTTTCGCTAACA	2040						
QY	3320	AAAAACATGTAAGTACAGTATCCAAAGAGGGAATATCATCTGAAGTATAGACAA	3379						
DB	2041	AAAAACATGTAAGTACAGTATCCAAAGAGGGAATATCATCTGAAGTATAGACAA	2100						
QY	3380	GCAACAGGCGAATCAGAACACTTACGGAGCCAGTGTGATTTAATCATAGTATGATTT	3439						

Db 2101 GCACAGCGGAATCAGAACACTTACGAGGCCAGTGTGATTTTAAATCATAGTGAAT 2160
Qy 3440 TTACACCCATATCCACTGTGTACAGAAAAATACAAATTAGAGACTTTCATTTATGGGAATA 3499
Db 2161 TTACACCCATATCCACTGTGTACAGAAAAATACAAATTAGAGACTTTCATTTATGGGAATA 2220
Qy 3500 AGCATTGAAGACACTGGTAGTACCAAGCATTTGGAGAAATAGACTTTAAAAATTCACCA 3559
Db 2221 AGCATTGAAGACACTGGTAGTACCAAGCATTTGGAGAAATAGACTTTAAAAATTCACCA 2280
Qy 3560 AGAATTTTGTGTACAGAGAAATCAGAGAAATACAAAGCCGAGAGAGGTTAGTGTAGAAA 3619
Db 2281 AGAATTTTGTGTACAGAGAAATCAGAGAAATACAAAGCCGAGAGAGGTTAGTGTAGAAA 2340
Qy 3620 GTTCAACGAGCTCACATATGAAGTACGTAGCCAAAGTTTCAATACAGACACTACAA 3679
Db 2341 GTTCAACGAGCTCACATATGAAGTACGTAGCCAAAGTTTCAATACAGACACTACAA 2400
Qy 3680 GTGGCATTAAGTTCAATGAGCTCAAGTCTTCCAGAGACAGTAGGTGTAGTACAA 3739
Db 2401 GTGGCATTAAGTTCAATGAGCTCAAGTCTTCCAGAGACAGTAGGTGTAGTACAA 2460
Qy 3740 CTATGGACACAGACTGTGGAAGCATGAGTACTGTGTAAAGTACTAAAACTATTAAAGCAA 3799
Db 2461 CTATGGACACAGACTGTGGAAGCATGAGTACTGTGTAAAGTACTAAAACTATTAAAGCAA 2520
Qy 3800 GCCACTATTGACGCCACAGCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCC 3859
Db 2521 GCCACTATTGACGCCACAGCTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCC 2580
Qy 3860 TGGTGCTTCCAGGTTCTTCTCATGCTTCTTAGAGAGACAGTCCCTTAAAGCACCCCT 3919
Db 2581 TGGTGCTTCCAGGTTCTTCTCATGCTTCTTAGAGAGACAGTCCCTTAAAGCACCCCT 2640
Qy 3920 CTATTGTACAAATAAAGTCTACAGATTTCTAATCTTGTAGTACACAGTTCTAGAGATG 3979
Db 2641 CTATTGTACAAATAAAGTCTACAGATTTCTAATCTTGTAGTACACAGTTCTAGAGATG 2700
Qy 3980 CTTTGTGCTATGCTACACTGAAAAGACTACAGCAACAAAGATGCATCCATCTTATCTC 4039
Db 2701 CTTTGTGCTATGCTACACTGAAAAGACTACAGCAACAAAGATGCATCCATCTTATCTC 2760
Qy 4040 ACTCTGAAGCTTTGGCATCTCAGCAAAAGATGTCTATTATCTGATACCAATCACCATGA 4099
Db 2761 ACTCTGAAGCTTTGGCATCTCAGCAAAAGATGTCTATTATCTGATACCAATCACCATGA 2820
Qy 4100 AGGCCAACAGTTTGTAGTCCAGATTTACCAACAGCGTTTCATGAAGCCCTTAAGTTATG 4159
Db 2821 AGGCCAACAGTTTGTAGTCCAGATTTACCAACAGCGTTTCATGAAGCCCTTAAGTTATG 2880
Qy 4160 CATCATTAGATAAAGAGATTTATGTGCTCTATTAAATCAAAATACCTGCAACGATCTT 4219
Db 2881 CATCATTAGATAAAGAGATTTATGTGCTCTATTAAATCAAAATACCTGCAACGATCTT 2940
Qy 4220 CCTCAGTGCCTGCATGGTGTCCAGTGCCACATATGGGGTTCAGATGATPACATTTGGTC 4279
Db 2941 CCTCAGTGCCTGCATGGTGTCCAGTGCCACATATGGGGTTCAGATGATPACATTTGGTC 3000
Qy 4280 TTGCTCTCCCGTGGATATAATGATATATATATATATATATATATATATATATATATATAT 4339
Db 3001 TTGCTCTCCCGTGGATATAATGATATATATATATATATATATATATATATATATATATAT 3060
Qy 4340 CAAAAACATACACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4399
Db 3061 CAAAAACATACACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Qy 4400 GTCTTCCATCTGGAATCTGGAGGTTCTGTAAAAAATCTTTTCACTTGTCTACGACAGCAGA 4459
Db 3121 GTCTTCCATCTGGAATCTGGAGGTTCTGTAAAAAATCTTTTCACTTGTCTACGACAGCAGA 3180
Qy 4460 TGAGTCTTACGGAATAATGAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 4519
Db 3181 TGAGTCTTACGGAATAATGAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 3240

Qy 4520 CAGAAGACACTGGACTACAGGAACATACAGATGATAACTGCCTTTATTTGTGTCTGTATTG 4579
Db 3241 CAGAAGACACTGGACTACAGGAACATACAGATGATAACTGCCTTTATTTGTGTCTGTATTG 3300
Qy 4580 AAATCTTGGGTTTCCAGCCCAAGCAACAACTGAGTGCAATATGTAGTCAATTCAGACTTTC 4639
Db 3301 AAATCTTGGGTTTCCAGCCCAAGCAACAACTGAGTGCAATATGTAGTCAATTCAGACTTTC 3360
Qy 4640 AAGATATTCCATATCTTGATTTGGTGAGCAGACTATCCATAATCCCTTTGAAGTGGTTC 4699
Db 3361 AAGATATTCCATATCTTGATTTGGTGAGCAGACTATCCATAATCCCTTTGAAGTGGTTC 3420
Qy 4700 CCTCTAAGTTTTCGGGGATTTCTGCATGCAGTGTGGGTCTCTCAAGAAAGCTCAGCTA 4759
Db 3421 CCTCTAAGTTTTCGGGGATTTCTGCATGCAGTGTGGGTCTCTCAAGAAAGCTCAGCTA 3480
Qy 4760 GCAGACCAAAAGCAGAGAAATTTGTACTAGGTGTTAAAAAATTTCCAGATGATACACAA 4819
Db 3481 GCAGACCAAAAGCAGAGAAATTTGTACTAGGTGTTAAAAAATTTCCAGATGATACACAA 3540
Qy 4820 TGTCCGCTATACCTCTTCGCAAAAGTTCCTAAGATTAGTCAATTAATTTGAGTAGTTCCAG 4879
Db 3541 TGTCCGCTATACCTCTTCGCAAAAGTTCCTAAGATTAGTCAATTAATTTGAGTAGTTCCAG 3600
Qy 4880 TTTCAACTAAATGTATGAGACTGGGCTTTTAAACAATTAAGGAGAAGTATCCTCAAAACAT 4939
Db 3601 TTTCAACTAAATGTATGAGACTGGGCTTTTAAACAATTAAGGAGAAGTATCCTCAAAACAT 3660
Qy 4940 TTGATGACATATGCTTTTACTCTGAGGTTTCCCAATTTGCTGTGCACATTCAGACTCAGAC 4999
Db 3661 TTGATGACATATGCTTTTACTCTGAGGTTTCCCAATTTGCTGTGCACATTCAGACTCAGAC 3720
Qy 5000 TTCCGCTGCGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCATG 5059
Db 3721 TTCCGCTGCGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCATG 3780
Qy 5060 AAGAAGCAGAGGCTGTGTGGCAACACACCAAGCAACCAAGCAACCAAGCAACCAAGCAAC 5119
Db 3781 AAGAAGCAGAGGCTGTGTGGCAACACACCAAGCAACCAAGCAACCAAGCAACCAAGCAAC 3840
Qy 5120 AATCCTGA 5127
Db 3841 AATCCTGA 3848

RESULT 5

ADC30081
ID ADC30081 standard; cDNA; 2659 BP.

XX AC ADC30081;

XX AC ADC30081;

XX DT 18-DEC-2003 (first entry)

XX DE Human novel cDNA sequence, SEQ ID NO:163.

XX KW Human; diagnostic; drug screening; forensics; gene mapping;
XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX KW ulcers; osteoporosis; autoimmune disease; cancer;
XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
XX KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
XX KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX KW gene therapy; chromosome 5; gene; ss.

OS Homo sapiens.

XX WO2003029271-A2.

XX PD 10-APR-2003.

XX PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
DR P-PSDB; ADC31052.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
XX Claim 1; SEQ ID NO 163; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2659 BP; 848 A; 539 C; 524 G; 748 T; 0 U; 0 Other;
Query Match 39.5%; Score 2026.4; DB 10; Length 2659;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 6; Indels 96; Gaps 1;
QY 2890 ATCAGAGGGACCTGTGTATATGACTTGGGCTCATAGCTAAACCAAGGCTGTGAT 2949
DB 268 AACCAAGGACCTGTGTATATGACTTGGGCTCATAGCTAAACCAAGGCTGTGAT 327
QY 2950 ATTCTAAATGTCAACCTGGATGCTGAGGATAGTCGAAACATCTGTGCGCAGTG 3009
DB 328 ATTCTAAATGTCAACCTGGATGCTGAGGATAGTCGAAACATCTGTGCGCAGTG 387
QY 3010 GTTCCAGATGATGGGAACAACTCTGTAATGAACCTTTTCATCTATCCCAAGCACTCTAAGT 3069
DB 388 GTTCCAGATGATGGGAACAACTCTGTAATGAACCTTTTCATCTATCCCAAGCACTCTAAGT 447
QY 3070 TTGAACCTGGAGTCAACCAAGCTCTAGACATAATAGTGAAGTGAATCTGTGCCATCGAGT 3129
DB 448 TTGAACCTGGAGTCAACCAAGCTCTAGACATAATAGTGAAGTGAATCTGTGCCATCGAGT 507

QY 3130 ATGTTTCATATTGGAGATGACCGGTTGGCAGCAGCTCTACTAGTACATATTTTCTCTTGAT 3189
DB 508 ATGTTTCATATTGGAGATGACCGGTTGGCAGCAGCTCTACTAGCACAATTTTCTCTTGAT 567
QY 3190 ATCAATGAAGATACAGAGCCCAACATTTTATGACCGATCTGACCCATTAAGAGATAAATAAT 3249
DB 568 ATCAATGAAGATACAGAGCCCAACATTTTATGACCGATCTGACCCATTAAGAGATAAATAAT 627
QY 3250 TCATTCCTCTTTCTTTGCTTCTAGTAAACTCTGTAAGATCGTATCTTAAATTCGCTTACT 3309
DB 628 TCATTCCTCTTTCTTTGCTTCTAGTAAACTCTGTAAGATCGTATCTTAAATTCGCTTACT 687
QY 3310 TTGCTTAACAAAACATCTGTAGTAGCAGTATCCAAAAGGAGGAAATATTCATCTGAA 3369
DB 688 TTGCTTAACAAAACATCTGTAGTAGCAGTATCCAAAAGGAGGAAATATTCATCTGAA 747
QY 3370 AGTAAGACAAGCAACAGCGGAATCAGAAACATTTACGGAGCCGAGTGTGATTTTAAATCAT 3429
DB 748 AGTAAGACAAGCAACAGCGGAATCAGAAACATTTACGGAGCCGAGTGTGATTTTAAATCAT 807
QY 3430 AGTGATGATTTTACACCCATATCCACTGTACAGAAACATTTACAAATTAGAGACTTCATTT 3489
DB 808 AGTGATGATTTTACACCCATATCCACTGTACAGAAACATTTACAAATTAGAGACTTCATTT 867
QY 3490 ATGGGGAATAAGCAATTTGAAGACACTCTGCTAGTACACCAAGCATTTGGAGAAATGACTTA 3549
DB 868 ATGGGGAATAAGCAATTTGAAGACACTCTGCTAGTACACCAAGCATTTGGAGAAATGACTTA 927
QY 3550 AAATTCACAAGAAATTTGGTACAGAGATCACAGAGAAATACAAAGCGAGAGAGTTA 3609
DB 928 AAATTCACAAGAAATTTGGTACAGAGATCACAGAGAAATACAAAGCGAGAGAGTTA 987
QY 3610 GTAGTAGAAGTTCAACGAGCTCACATATGAAGTACCTAGTCCAAAGTTTCAATACAGAC 3669
DB 988 GTAGTAGAAGTTCAACGAGCTCACATATGAAGTACCTAGTCCAAAGTTTCAATACAGAC 1047
QY 3670 ACTACAACAAGTGGCATTAAGTTCAATGAGCTCAAGTCTCTTACAGAGACAGTAGGTGTA 3729
DB 1048 ACTACAACAAGTGGCATTAAGTTCAATGAGCTCAAGTCTCTTACAGAGACAGTAGGTGTA 1107
QY 3730 GATGTACAACATATGGACACAGACTGTGGAAGCATGAGTACTGTGTAAGTACTAAACT 3789
DB 1108 GATGTACAACATATGGACACAGACTGTGGAAGCATGAGTACTGTGTAAGTACTAAACT 1167
QY 3790 ATTAAGACAAGCCACTTATTTGACGCGCACAGTCTAACCACTCTCTCTCTCCAAATCAAT 3849
DB 1168 ATTAAGACAAGCCACTTATTTGACGCGCACAGTCTAACCACTCTCTCTCTCCAAATCAAT 1227
QY 3850 TCGGTGTCCCTGGTGCCTCCAGGTTCTTCTCATACGCTTCTTAGAAGACACAGTCCCTT 3909
DB 1228 TCGGTGTCCCTGGTGCCTCCAGGTTCTTCTCATACGCTTCTTAGAAGACACAGTCCCTT 1287
QY 3910 AAAGCACCTCTATTGTCTACATTAAGTCTAGCAGATTTGTAACCTTTAGTTTACACAAGT 3969
DB 1288 AAAGCACCTCTATTGTCTACATTAAGTCTAGCAGATTTGTAACCTTTAGTTTACACAAGT 1347
QY 3970 TCTAGAGATGCTTTTGGCTTATGCTACCTGAAAGACTACAGCAACAAAGAAATGATCCA 4029
DB 1348 TCTAGAGATGCTTTTGGCTTATGCTACCTGAAAGACTACAGCAACAAAGAAATGATCCA 1407
QY 4030 TCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGTATCC 4089
DB 1408 TCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTTACTGTATCC 1467
QY 4090 ATCACCATGAAGCCCAACAGTTTTCAGTCCAGATTAACCAAGCAGGTTTCATGAAAGCC 4149
DB 1468 ATCACCATGAAGCCCAACAGTTTTCAGTCCAGATTAACCAAGCAGGTTTCATGAAAGCC 1527
QY 4150 TTAAGTTATGTCATCAATTAGATAAAGAGATTTTATGAGTCTTATTAATCAAAATACCCCTG 4209
DB 1528 TTAAGTTATGTCATCAATTAGATAAAGAGATTTTATGAGTCTTATTAATCAAAATACCCCTG 1587
QY 4210 CAACGATCTTCTCAGTGGGCTCCATGCTGCCAGTCCACATATGGGGGTTTCAGATGAT 4269

Db 1588 CAACGATCTTCTCAGTCGGTCCATGTCAGTGCACATATGGGGTTCAGATGAT 1647
Qy 4270 TACATTGCTCTGCTCTCCGGTGGATATAATGATATATTCAGGTAAAGATATTCCTC 4329
Db 1648 TACATTGCTCTGCTCTCCGGTGGATATAATGATATATTCAGGTAAAGATATTCCTC 1707
Qy 4330 TATTTCAGCAAAAAACATACCAACATGATGATGCGAGTGCAGAGCATTTTGGCCAT 4389
Db 1708 TATTTCAGCAAAAAACATACCAACATGATGATGCGAGTGCAGAGCATTTTGGCCAT 1767
Qy 4390 GATCAGGAGGTCTTCATCTGGAACCTGGAGTCTTGTAAAAATCTTTTCACTTGCTA 4449
Db 1768 GATCAGG----- 1775
Qy 4450 CGACAGCAGATGAGTCTTACGGAATAATGAATCCATTCAGATGCGCTCTCTGTTT 4509
Db 1776 -----AGATGCTCTCTGTTT 1791
Qy 4510 TTAGAAAGTACAGAGACACTGGACTACAGGAACATACAGATGATGAATGCTTTATGT 4569
Db 1792 TTAGAAAGTACAGAGACACTGGACTACAGGAACATACAGATGATGAATGCTTTATGT 1851
Qy 4570 GTCTGATTGAATCTGGTTTCAGCCAGCAACCACTGAGTGCATATGATGATCAT 4629
Db 1852 GTCTGATTGAATCTGGTTTCAGCCAGCAACCACTGAGTGCATATGATGATCAT 1911
Qy 4630 TCAGACTTTCAAGATATTCATATTTCTGATGTTGTGAGCAGATTCATCAATCCCTTA 4689
Db 1912 TCAGACTTTCAAGATATTCATATTTCTGATGTTGTGAGCAGATTCATCAATCCCTTA 1971
Qy 4690 GAAGTGTTCCTCTAAGTTTTCGGGATTTCTGGATGACATGATGGGTGTCTCAAGAA 4749
Db 1972 GAAGGGTTCCTCTAAGTTTTCGGGATTTCTGGATGACATGATGGGTGTCTCAAGAA 2031
Qy 4750 GGCTCAGCTAGCAGCAACAAAGCAGAAATTTGTTACTAGGTGTTAAACAAATTCAGAT 4809
Db 2032 GGCTCAGCTAGCAGCAACAAAGCAGAAATTTGTTACTAGGTGTTAAACAAATTCAGAT 2091
Qy 4810 GATACCAATGTCGGTATCTCTTCGAAAGAGTCTTAAGATGATCAATTAATTTG 4869
Db 2092 GATACCAATGTCGGTATCTCTTCGAAAGAGTCTTAAGATGATCAATTAATTTG 2151
Qy 4870 AGTAGTTTCAGTTTCAACTAAATGTCATGAGACTGGGCTTTTAAACAAATTAAGGAGAT 4929
Db 2152 AGTAGTTTCAGTTTCAACTAAATGTCATGAGACTGGGCTTTTAAACAAATTAAGGAGAT 2211
Qy 4930 CCTCAACATTTGATGATGATGCTTTTACTCTGAGGTTTCCCATTTGCTGTCACTGC 4989
Db 2212 CCTCAACATTTGATGATGATGCTTTTACTCTGAGGTTTCCCATTTGCTGTCACTGC 2271
Qy 4990 ACATTGACATTCGTTGCGAGTTCATACAGAAATTTTCAAGATGTACAGTTTCTA 5049
Db 2272 ACATTGACATTCGTTGCGAGTTCATACAGAAATTTTCAAGATGTACAGTTTCTA 2331
Qy 5050 CAAATGATGAAGAGCAGAGGCTGTGTGCAACACCCAAACCAACCTATAGTTGAT 5109
Db 2332 CAAATGATGAAGAGCAGAGGCTGTGTGCAACACCCAAACCAACCTATAGTTGAT 2391
Qy 5110 ACATCTGCTGAATCCTGA 5127
Db 2392 ACATCTGCTGAATCCTGA 2409

RESULT 6
ADL45178/c
ID ADL45178 standard; DNA; 4657 BP.
XX
AC ADL45178;
XX
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #19068.

XX
KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX OS Homo sapiens.
XX WO200170979-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009126.
XX PR 21-MAR-2000; 2000US-0191031P.
XX PR 25-MAY-2000; 2000US-0207124P.
XX PR 15-JUN-2000; 2000US-0211940P.
XX PR 07-JUL-2000; 2000US-0216820P.
XX PR 25-JUL-2000; 2000US-0220661P.
XX PR 21-DEC-2000; 2000US-0257672P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lee J, Lillie J;
XX WPI; 2001-611502/70.
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
XX Disclosure; SEQ ID NO 19068; 106pp; English.
XX The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention.
XX
SQ Sequence 4657 BP; 1440 A; 860 C; 877 G; 1476 T; 0 U; 4 Other;
Query Match 34.6%; Score 1772.8; DB 5; Length 4657;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 2; Indels 97; Gaps 1;
Qy 3148 GACCGGTTTGGCAGCAGCTCTACTATACATTTTCTCTCATATCATGATACAGAG 3207
Db 4634 GACCGGTTTGGCAGCAGCTCTACTATACATTTTCTCTCATATCATGATACAGAG 4575
Qy 3208 CCAACATTTTATGACCGATCTGGACCCCAAGATAAAAAATTCATTCTTCTTCTTCT 3267

Db	4574	CCAA	CACTTTTATGACCGGACTCGACCCATAAAGGATATAAAATTCATTCCCTTTCTTTGGCT	4515
Qy	3268	TTCTAGTAAAC	CTTGTGGAAGAAATCGTATCTTAAATTCGCTTACTTTTGCCCTAACAAAAACAT	3327
Db	4514	TTCTAGTAAAC	CTTGTGGAAGAAATCGTATCTTAAATTCGCTTACTTTTGCCCTAACAAAAACAT	4455
Qy	3328	CGTAGTAGCAGTGAT	CCAAAAAGGAGGAAATTTATCATCTGAAAGTAAAGACAAGCAACAGG	3387
Db	4454	CGTAGTAGCAGTGAT	CCAAAAAGGAGGAAATTTATCATCTGAAAGTAAAGACAAGCAACAGG	4395
Qy	3388	CGAATCAGAACACTT	ACGAGGCCAGTGTGATTTTAAATCATATAGTAGTATGATTTTACACCC	3447
Db	4394	CGAATCAGAACACTT	ACGAGGCCAGTGTGATTTTAAATCATATAGTAGTATGATTTTACACCC	4335
Qy	3448	ATATCCACTGTGACAGAAAA	CACTTACAAATTTAGAGACTTCATTTATCGGGGAATAAGCACAT	3507
Db	4334	ATATCCACTGTGACAGAAAA	CACTTACAAATTTAGAGACTTCATTTATCGGGGAATAAGCACAT	4275
Qy	3508	GAAGACACTCGTGTAGTAC	CCACAGCAATTTGAGAAAAATGACTTAAAAATTCACAGAAATTTT	3567
Db	4274	GAAGACACTCGTGTAGTAC	CCACAGCAATTTGAGAAAAATGACTTAAAAATTCACAGAAATTTT	4215
Qy	3568	GGTACAGAGAAATCAC	AGAGAAAAATACACGCCAGAGAGAGTGTAGTAGTAGAAAAAGTTCAACG	3627
Db	4214	GGTACAGAGAAATCAC	AGAGAAAAATACACGCCAGAGAGAGTGTAGTAGTAGAAAAAGTTCAACG	4155
Qy	3628	AGCTCACATATGAAGATAC	GTAGCCCAAAGTTTCAATACAGACACTACAAACAAGTGGCATA	3687
Db	4154	AGCTCACATATGAAGATAC	GTAGCCCAAAGTTTCAATACAGACACTACAAACAAGTGGCATA	4095
Qy	3688	AGTTCAATAGACTCAAGT	CTCTTCCAGAGACAGTAGTAGTGCTACTCAACTATGGAC	3747
Db	4094	AGTTCAATAGACTCAAGT	CTCTTCCAGAGACAGTAGTAGTGCTACTCAACTATGGAC	4035
Qy	3748	ACAGACTGTGGAAGCATG	ACTGCTGGTAAAGTACTTAAACCTAATTAAAGACAAGCCACTAT	3807
Db	4034	ACAGACTGTGGAAGCATG	ACTGCTGGTAAAGTACTTAAACCTAATTAAAGACAAGCCACTAT	4009
Qy	3808	TTGAGCCACAGTCTAAC	CACTCTCTCTCCAAATCAAATTCGGTGTCCCTGGTGCCCT	3867
Db	4008	-----	-----	4009
Qy	3868	CCAGGTTCTTCTCATAC	GCCTTCTTAGAAGAGCACAGTCCCTTAAAGCACCCCTCTATTGCT	3927
Db	4008	---GGTCTTCTCATAGC	TCTCTTAGAAGAGCACAGTCCCTTAAAGCACCCCTCTATTGCT	3952
Qy	3928	ACAAATAAAGTCTAGCAG	ATTGTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGC	3987
Db	3951	ACAAATAAAGTCTAGCAG	ATTGTAACTTTAGTTTACACAAGTTCTAGAGATGCTTTTGGC	3892
Qy	3988	TATGCTACACTGMAAAG	ACTACAGCAAAAGAAATGCATCCATCCTTTATCTCACTGTAA	4047
Db	3891	TATGCTACACTGMAAAG	ACTACAGCAAAAGAAATGCATCCATCCTTTATCTCACTGTAA	3832
Qy	4048	GCTTTTGGCATCTCCAG	CAAAAAGATGTGTATTATCTGATACCATCACATGAGGCGCAAC	4107
Db	3831	GCTTTTGGCATCTCCAG	CAAAAAGATGTGTATTATCTGATACCATCACATGAGGCGCAAC	3772
Qy	4108	AGTTTTTGAGTCAGATTA	ACAACCAAGAGGTTCTATGAAAGCCTTTAAGTTATGATCATTA	4167
Db	3771	AGTTTTTGAGTCAGATTA	ACAACCAAGAGGTTCTATGAAAGCCTTTAAGTTATGATCATTA	3712
Qy	4168	GATAAAGAAAGATTTAT	TGAGTCCCTATTAAATCAAAATACCCCTGGCAACGATCTTCCCTCAGTG	4227
Db	3711	GATAAAGAAAGATTTAT	TGAGTCCCTATTAAATCAAAATACCCCTGGCAACGATCTTCCCTCAGTG	3652
Qy	4228	CGGTCCATCGGTCCAGT	CGCCACATATGGGGTTCAGATGATTAATATGGTCTTGGCTCTC	4287
Db	3651	CGGTCCATCGGTCTCCAG	TGCCACATATGGGGTTCAGATGATTAATATGGTCTTGGCTCTC	3592
Qy	4288	CCGGTGGATATAAATG	ATATATTTCCAGGTAAAGGATATTTCCCTATTTTTCAGACAAAAAC	4347

Db	3591	CCGGTGGATATAAATCGATATATTTCCAGGTAAAGGATATTCCTTATTTTTCAGCAAAAAC	3532
Qy	4348	ATACCAACCATGATGATCGAGGTGCAAGAGCATTTGCCCATGATGCAGGAGGTCTTCCA	4407
Db	3531	ATACCAACCATGATGATCGAGGTGCAAGAGCATTTGCCCATGATGCAGGAGGTCTTCCA	3472
Qy	4408	TCTGGAACTGGAGGTCTTCTGTAATAATCTTTTCACTTGTACGACACAGATGAGTCTT	4467
Db	3471	TCTGGAACTGGAGGTCTTGTGTAATAATCTTTTCACTTGTACGACACAGATGAGTCTT	3412
Qy	4468	ACGGAAATAATGAATTCATTCACATTCAGATGCCTCTCTGTTTTTATGAAAGTACAGAAGAC	4527
Db	3411	ACGGAAATAATGAATTCATTCACATTCAGATGCCTCTCTGTTTTTATGAAAGTACAGAAGAC	3352
Qy	4528	ACTGGACTACAGAAACATACAGATGATTAAGTGCCTTTATTTGTTCTGTATTTGAAATCTTG	4587
Db	3351	ACTGGACTACAGAAACATACAGATGATTAAGTGCCTTTATTTGTTCTGTATTTGAAATCTTG	3292
Qy	4588	GGTTTTCCAGCCACGACCAACCACTGAGTGCATAATGTAGTCAATTCAGACTTTTCAAGATATT	4647
Db	3291	GGTTTTCCAGCCACGACCAACCACTGAGTGCATAATGTAGTCAATTCAGACTTTTCAAGATATT	3232
Qy	4648	CCATATTCTGATTTGGTGTGAGCAGACTATCCATPAATCCCTTTAGAAAGTGTTCCCTCTAAG	4707
Db	3231	CCATATTCTGATTTGGTGTGAGCAGACTATCCATPAATCCCTTTAGAAAGTGTTCCCTCTAAG	3172
Qy	4708	TTTTCCGGGATTTCTCGATCGAGTATGGGTGTCTCAAGAGGCTCAGCTAGCAGCAC	4767
Db	3171	TTTTCCGGGATTTCTCGATCGAGTATGGGTGTCTCAAGAGGCTCAGCTAGCAGCAC	3112
Qy	4768	AAAAGCACAGAAATTTTACTAGGTGTTTAAACCAATTCAGATGATACACCAATGTGCCGT	4827
Db	3111	AAAAGCACAGAAATTTTACTAGGTGTTTAAACCAATTCAGATGATACACCAATGTGCCGT	3052
Qy	4828	ATACTCTTCCGAAAGAAGTCTTAAGATTAGTCAATTAATTTGATGATTTCACTTTCAACT	4887
Db	3051	ATACTCTTCCGAGAGAAGTCTTAAGATTAGTCAATTAATTTGATGATTTCACTTTCAACT	2992
Qy	4888	AAATGCTATGAGACTGGGCTTTTAAACCAATTAAGGAGAAGTATCCTCAACCAATTTGATGAC	4947
Db	2991	AAATGCTATGAGACTGGGCTTTTAAACCAATTAAGGAGAAGTATCCTCAACCAATTTGATGAC	2932
Qy	4948	ATATGCTTTTACTCTGAGGTTTTCCCATTTGCTGTGCACACTGCACATTCAGACTTCGGTGT	5007
Db	2931	ATATGCTTTTACTCTGAGGTTTTCCCATTTGCTGTGCACACTGCACATTCAGACTTCGGTGT	2872
Qy	5008	CGGAGGTTTCATACAAAGAAATTTTCAAGATGTAAGTTTTCTAATAATGCAATGAGAGCA	5067
Db	2871	CGGAGGTTTCATACAAAGAAATTTTCAAGATGTAAGTTTTCTAATAATGCAATGAGAGCA	2812
Qy	5068	GAGGCTGTGTTGGCAACACCAAGCAACCTATAGTTGATACATCTGCTGAATCTCTGA	5127
Db	2811	GAGGCTGTGTTGGCAACACCAAGCAACCTATAGTTGATACATCTGCTGAATCTCTGA	2752

RESULT 7
ABV20190/c
in ABV20190 standard: cDNA: 4601 BP.

XX
AC
XX
DT
XX
DE

ABV20190;

13-SEP-2002 (first entry)

Human prostate expression

human prostate expression marker CDNA 20181.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

Homo sapiens.

WC200160860-A2.

23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.
XX PF
XX PR
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3296; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 4601 BP; 1416 A; 852 C; 871 G; 1460 T; 0 U; 2 Other;

Query Match 34.5%; Score 1768; DB 5; Length 4601;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 5; Indels 97; Gaps 1;

QY 3148 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCCTTGATATCAATGAAGATACAGAG 3207
DB 4580 GACCGGTTGGCAGCAGCTCTACTAGTACATTTTCCTTGATATCAATGAAGATACAGAG 4521
QY 3208 CCAACATTTTATGACCGATCTGGACCCCAATAAGGATATAATTCATTCCTTTCTTTGCT 3267
DB 4520 CCAACATTTTATGACCGATCTGGACCCCAATAAGGATATAATTCATTCCTTTCTTTGCT 4461
QY 3268 TCTAGTAACTTGTAAGAATCGTATCTTAAATTCGCTTACTTTTGCCTTAAACAAAAACAT 3327
DB 4460 TCTAGTAACTTGTAAGAATCGTATCTTAAATTCGCTTACTTTTGCCTTAAACAAAAACAT 4401
QY 3328 CGTAGTAGCAGTATCCAAAGGAGGGGAAATATCATCTGAAATGACAGCAAGCAAGG 3387
DB 4400 CGTAGTAGCAGTATCCAAAGGAGGGGAAATATCATCTGAAATGACAGCAAGCAAGG 4341
QY 3388 CGAATCAGAACATTTACGGAGCCAGTGTTGATTTTAAATCATAGTATGATTTTACACCC 3447
DB 4340 CGAATCAGAACATTTACGGAGCCAGTGTTGATTTTAAATCATAGTATGATTTTACACCC 4281
QY 3448 ATATCCACTGTACAGAAAAACATTACAATTAAGAGACTTCATTTATGGGGAATAAGCAAT 3507
DB 4280 ATATCCACTGTACAGAAAAACATTACAATTAAGAGACTTCATTTATGGGGAATAAGCAAT 4221
QY 3508 GAAGACATGTTAGTACACCAAGCATTTGGAGAAAAATGACTTTAAATTTACCAAGAAATTT 3567
DB 4220 GAAGACATGTTAGTACACCAAGCATTTGGAGAAAAATGACTTTAAATTTACCAAGAAATTT 4161
QY 3568 GGTACAGAGATCACAGAGAAAAATACAGCCGAGAGGTTAGTAGTAAGTTCAACG 3627
DB 4160 GGTACAGAGATCACAGAGAAAAATACAGCCGAGAGGTTAGTAGTAAGTTCAACG 4101

QY 3628 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACCAAGTGGCATA 3687
DB 4100 AGCTCACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACCAAGTGGCATA 4041
QY 3688 AGTTCAATGAGCTCAAGTCCCTTCCAGGAGAGACAGTAGTGTAGTGTCTACAACATATGAC 3747
DB 4040 AGTTCAATGAGCTCAAGTCCCTTCCAGGAGAGACAGTAGTGTAGTGTCTACAACATATGAC 3981
QY 3748 ACAGACTGTGGAAGCATGAGTACTGTGTAAGTACTAAAACTATTAAGACAAGCCACTAT 3807
DB 3980 ACAGACTGTGGAAGCATGAGTACTGT----- 3955
QY 3808 TTGAGGCCACAGTCTAACCATCTGTCTCTCTCCAAATCAAAATTCGGTGTCTCTGGTGCCT 3867
DB 3954 ----- 3955
QY 3868 CCAGGTTCTTCTCATACGCTTCCCTAGAGACAGACAGTCCCTTAAAGCAACCCCTCTATTGCT 3927
DB 3954 ---GGTCTCTCCGTCAGGCTTCCCTAGAGAGACAGTCCCTTAAAGCAACCCCTCTATTGCT 3988
QY 3928 ACAATTAAAGTCTAGCAGATTTGTAAGTTTGTAGTTACACAAGTCTCTAGAGATGCTTTTGGC 3987
DB 3997 ACAATTAAAGTCTAGCAGATTTGTAAGTTTGTAGTTACACAAGTCTCTAGAGATGCTTTTGGC 3838
QY 3988 TATGCTACACTGAAAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTGAA 4047
DB 3837 TATGCTACACTGAAAAAGACTACAGCAACAAAGAAATGCATCCATCTTATCTCACTGAA 3778
QY 4048 GCTTTGCACTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGAAGGCCAAC 4107
DB 3777 GCTTTGCACTCCAGCAAAAGATGTGCTATTTACTGATACCATCACCATGAAGGCCAAC 3718
QY 4108 AGTTTTCAGTCCCATTAACACCAAGCAGGTTCAATGAAAGCTTAAAGTTATGTCATCTTA 4167
DB 3717 AGTTTTCAGTCCCATTAACACCAAGCAGGTTCAATGAAAGCTTAAAGTTATGTCATCTTA 3658
QY 4168 GATAAAGAAGATTTATGAGTCTTATTAATCAAAATACCCCTGCAACGATCTTCTCAGTG 4227
DB 3657 GATAAAGAAGATTTATTTAGTCTTATTAATCAAAATACCCCTGCAACGATCTTCTCAGTG 3598
QY 4228 CGGTCCATGTGTCCAGTGCCACATATGCGGGTTCCAGATGATTAATGCTTGTCTCTC 4287
DB 3597 CGGTCCATGTGTCCAGTGCCACATATGCGGGTTCCAGATGATTAATGCTTGTCTCTC 3538
QY 4288 CCGGTGATATAAATGATATATCCAGTAAAGATATTCCTCTATTTTTCAGCAAAAAAAC 4347
DB 3537 CCGGTGATATAAATGATATATCCAGTAAAGATATTCCTCTATTTTTCAGCAAAAAAAC 3478
QY 4348 ATACCAACCATGATGATGAGGTGCAAGAGCATTTGCCCATGATGCAAGGAGTCTTCCA 4407
DB 3477 ATACCAACCATGATGATGAGGTGCAAGAGCATTTGCCCATGATGCAAGGAGTCTTCCA 3418
QY 4408 TCTGGAATGAGGTCTTGTAAAAAATTTCTTTTCACTTGTCTACGACAGCAGATGATCTT 4467
DB 3417 TCTGGAATGAGGTCTTGTAAAAAATTTCTTTTCACTTGTCTACGACAGCAGATGATCTT 3358
QY 4468 ACGGAATAATGATTCATTCATTCAGATGCCCTCTCTGTTTTTAAAGAAATGACAGAAC 4527
DB 3357 ACGGAATAATGATTCATTCATTCAGATGCCCTCTCTGTTTTTAAAGAAATGACAGAAC 3298
QY 4528 ACTGGACTACAGGAACATACAGATGATACTGCCCTTTATTTGTGTCTGTATTGAAATCTG 4587
DB 3297 ACTGGACTACAGGAACATACAGATGATACTGCCCTTTATTTGTGTCTGTATTGAAATCTG 3238
QY 4588 GGTTCAGGCCCAAGAACCAACTGAGTGCAATATGATGATCAATTCAGACTTTTCAAGATATT 4647
DB 3237 GGTTCAGGCCCAAGAACCAACTGAGTGCAATATGATGATCAATTCAGACTTTTCAAGATATT 3178
QY 4648 CCATATTTCTGATTTGTTGAGCAGACTATCCATAATCTTTTAGAAGTGGTTCCCTCTTAAG 4707
DB 3177 CCATATTTCTGATTTGTTGAGCAGACTATCCATAATCTTTTAGAAGTGGTTCCCTCTTAAG 3118
QY 4708 TTTTCGGGATTTCTGATGATGATGAGGTGTTCTCAAGAAAGGCTCAGCTAGCAGCACC 4767

Db 3117 TTTTCGGGATTTCTGGATGCGAGTATGGGGTCTCAAGAGGCTCAGTAGCAGACC 3058
Qy 4768 AAAAGCACAGATTTCTAGGTGTTAAACAAATTCAGATGATACACCAATGTGCCGT 4827
Db 3057 AAAAGCACAGATTTCTAGGTGTTAAACAAATTCAGATGATACACCAATGTGCCGT 2998
Qy 4828 ATATCTCTTCGCAAGAGTTCTAAGATTAGTATCATTAATTTGAGTAGTTCAGTTCAACT 4887
Db 2997 ATATCTCTTCGCAAGAGTTCTAAGATTAGTATCATTAATTTGAGTAGTTCAGTTCAACT 2938
Qy 4888 AATGTCATGACATGGGCTTTTAAACAATTAAGGAGAGTATCTCTCAACATTTGATGAC 4947
Db 2937 AATGTCATGACATGGGCTTTTAAACAATTAAGGAGAGTATCTCTCAACATTTGATGAC 2878
Qy 4948 ATATGCTTTTACTCTGAGGTTTCCCATTTGCTGTCACTGCACATTCAGACTCCGCTGT 5007
Db 2877 ATATGCTTTTACTCTGAGGTTTCCCATTTGCTGTCACTGCACATTCAGACTCCGCTGT 2818
Qy 5008 CGGAGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCAATGCAAGAGCA 5067
Db 2817 CGGAGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTACAAATGCAATGCAAGAGCA 2758
Qy 5068 GAGGCTGTGTGCAACACACCAAGCAACTATAGTTGATACATCTGCTGAATCCTGA 5127
Db 2757 GAGGCTGTGTGCAACACACCAAGCAACTATAGTTGATACATCTGCTGAATCCTGA 2698

RESULT 8
ABV26019/c
ID ABV26019 standard; cDNA; 4602 BP.
XX AC ABV26019;
DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 26010.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; SB.
XX OS Homo sapiens.
XX PN W0200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 5261; 11750pp; English.
XX PS The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy

of a test compound to inhibit prostate cancer in a patient; (d) assessing
the efficacy of a therapy for inhibiting prostate cancer in a patient;
(e) selecting a composition for inhibiting prostate cancer in a patient;
(f) assessing the prostate cell carcinogenic potential of a compound; (g)
determining whether prostate cancer has metastasized in a patient; (h)
assessing the aggressiveness or indolence of prostate cancer in a patient
; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX Sequence 4602 BP; 1416 A; 852 C; 871 G; 1460 T; 0 U; 3 Other;
Query Match 34.5%; Score 1768; DB 5; Length 4602;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 5; Indels 97; Gaps 1;
Qy 3148 GACCGTTTGGCAGCAGCTCTACTAGTACATTTTTTCTTGATATCAATGAAGATACAGAG 3207
Db 4580 GACCGTTTGGCAGCAGCTCTACTAGTACATTTTTTCTTGATATCAATGAAGATACAGAG 4521
Qy 3208 CCAACATTTTATGACCGATCTGGACCCCATAAAGGATAAAAATTCATTCCTTTCTTGCT 3267
Db 4520 CCAACATTTTATGACCGATCTGGACCCCATAAAGGATAAAAATTCATTCCTTTCTTGCT 4461
Qy 3268 TCTAGTAACTTTGAGAGATCGTATCTTAAATTCGCTTACTTTGCCTAAACAAAACAT 3327
Db 4460 TCTAGTAACTTTGAGAGATCGTATCTTAAATTCGCTTACTTTGCCTAAACAAAACAT 4401
Qy 3328 CGTAGTAGCAGTATCCAAAAGGAGGAAATTTATCATCTGAAAGTAAGCAAGCAACAGG 3387
Db 4400 CGTAGTAGCAGTATCCAAAAGGAGGAAATTTATCATCTGAAAGTAAGCAAGCAACAGG 4341
Qy 3388 CGAATCAGAACACTTTACGGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACC 3447
Db 4340 CGAATCAGAACACTTTACGGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACC 4281
Qy 3448 ATATCCACTGTACAGAAAACATTAATTCAGTACTTCATTTATGGGAAATTAAGCACAT 3507
Db 4280 ATATCCACTGTACAGAAAACATTAATTCAGTACTTCATTTATGGGAAATTAAGCACAT 4221
Qy 3508 GAAGACACTGTGTAGTACACCAAGCATTTGGAGAAATGACTTTAAATTCACCAAGAAATTT 3567
Db 4220 GAAGACACTGTGTAGTACACCAAGCATTTGGAGAAATGACTTTAAATTCACCAAGAAATTT 4161
Qy 3568 GGTACAGAGATTCACAGAGAAAATCAAGCCGAGAGAGTTAGTATGAGTAAAGTTCAACG 3627
Db 4160 GGTACAGAGATTCACAGAGAAAATCAAGCCGAGAGAGTTAGTATGAGTAAAGTTCAACG 4101
Qy 3628 AGCTCACATATGAAGTACGTAGCCAAAGTTTCAATACAGACACTACCAAGTGGCATA 3687
Db 4100 AGCTCACATATGAAGTACGTAGCCAAAGTTTCAATACAGACACTACCAAGTGGCATA 4041
Qy 3688 AGTTCAATGAGCTCAAGTCTTTACGAGAGACAGTAGGTGATGCTTACAACTATGGAC 3747
Db 4040 AGTTCAATGAGCTCAAGTCTTTACGAGAGACAGTAGGTGATGCTTACAACTATGGAC 3981
Qy 3748 ACAGACTGTGGAAGCATGAGTACTGTGTAAGTAACTTAAAGCAAGCCACTAT 3807
Db 3980 ACAGACTGTGGAAGCATGAGTACTGTGTAAGTAACTTAAAGCAAGCCACTAT 3955
Qy 3808 TTGAGCCACAGCTAAACCATCTGCTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCT 3867
Db 3954 ----- 3955
Qy 3868 CCAGGTTCTTCTCATACGCTTCTAGAGAGACAGTCCCTTTAAAGCACCCCTCTATTGCT 3927
Db 3954 ---GGTTCTTCGTCAGGCTTCTAGAGAGACAGTCCCTTTAAAGCACCCCTCTATTGCT 3898
Qy 3928 ACATTTAAAGTCTAGCAGTGTAACTTTAGTTACACAAGTTCTAGAGATGCTTTTGGC 3987
Db 3897 ACATTTAAAGTCTAGCAGTGTAACTTTAGTTACACAAGTTCTAGAGATGCTTTTGGC 3838
Qy 3988 TATGCTACACTGAAAGAGCTACAGCAACAAAGATGCTATCCCTTATCTCACTCTGAA 4047
Db 3837 TATGCTACACTGAAAGAGCTACAGCAACAAAGATGCTATCCCTTATCTCACTCTGAA 3778

QY 181 AATAACTTTACTAAGCTTCTTTGTGATATTTGGCCACAGTGAAGAAAACTGGGCTTTCAC 240
 Db 191 AATACTTTACTAAGCTTCTTTGTGATATTTGGCCACAGTGAAGAAAACTGGGCTTTCAC 250
 QY 241 TAGAGGATATCATATTTGTTGGGTTAGCTTTATTAATGAAGCAAAAGAGTGGCA 300
 Db 251 TAGAGGATATCATATTTGTTGGGTTAGCTTTATTAATGAAGCAAAAGAGTGGCA 310
 QY 301 GCAGAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTTCTCAGAAAGGTG 360
 Db 311 GCAGAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTTCTCAGAAAGGTG 370
 QY 361 CTAAAAATGAAAGTGGACTATTTAATAGCTAGTGCATTCATCAACAGAGCAACGAG 420
 Db 371 CTAAAAATGAAAGTGGACTATTTAATAGCTAGTGCATTCATCAACAGAGCAACGAG 430
 QY 421 GTAGAGAGCACAGACCTTCGATAGTGCAGAAAGATGATGTAATCTTCCCTTG 480
 Db 431 GTAGAGAGCACAGACCTTCGATAGTGCAGAAAGATGATGTAATCTTCCCTTG 490
 QY 481 TTTCTAGTTCTGTGACCAACTCATTAATTTGAGTTGGAAATGATGACTTCAAGAAAGA 540
 Db 491 TTTCTAGTTCTGTGACCAACTCATTAATTTGAGTTGGAAATGATGACTTCAAGAAAGA 550
 QY 541 GACAGAAATGTCGAGATGCTGATATTCGTGAATAGTACCTTCAGATCCAGAG 600
 Db 551 GACAGAAATGTCGAGATGCTGATATTCGTGAATAGTACCTTCAGATCCAGAG 610
 QY 601 GTGGTGGCCCTTCGAGAGGACTTAAACACCATATTTGAAATGATGATTTGCCAATTA 660
 Db 611 GTGGTGGCCCTTCGAGAGGACTTAAACACCATATTTGAAATGATGATTTGCCAATTA 670
 QY 661 AGTCGAATAATGAGGCCCTTAATTTACTACAAATTTGCACTTTTAAATCATCCAAAGCT 720
 Db 671 AGTCGAATAATGAGGCCCTTAATTTACTACAAATTTGCACTTTTAAATCATCCAAAGCT 730
 QY 721 CGGAGTATGTCGAGCTGATAGTATGAGAGAAATTTAGCACCCTTACTGATTTT 780
 Db 731 CGGAGTATGTCGAGCTGATAGTATGAGAGAAATTTAGCACCCTTACTGATTTT 790
 QY 781 CACTACAGACATAGTCAGATACAGCTGAAGGACAGCTCAAGAAAGACAGAGACGCA 840
 Db 791 CACTACAGACATAGTCAGATACAGCTGAAGGACAGCTCAAGAAAGACAGAGACGCA 850
 QY 841 TTTCTAGCCAGTAAATGAGGATCATAGCAATTCGATCATCGGACGATATTAAT 900
 Db 851 TTTCTAGCCAGTAAATGAGGATCATAGCAATTCGATCATCGGACGATATTAAT 910
 QY 901 TTATGTAACCTGGAATTTCTGGATCCAGTCTCTAATAGGAGTACTTTGCATACCAAT 960
 Db 911 TTATGTAACCTGGAATTTCTGGATCCAGTCTCTAATAGGAGTACTTTGCATACCAAT 970
 QY 961 ATGGAATAAGCGAGGTCTACTTTGAAGTGCCTTTATGATATATTTGCTTT 1011
 Db 971 ATGGAATAAGCGGTAGGTAAAAGTTATAGGATTTGTTTCTCTCGACCT 1021

RESULT 10
 ADR08103
 ID ADR08103 standard; cDNA; 3129 BP.
 AC ADR08103;
 XX
 XX
 DT 04-NOV-2004 (first entry)
 DE
 DE
 KW Full length human cDNA useful for treating neurological disease Seq 1609.
 KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
 KW osteoporosis; neurological disease; Alzheimer's disease;
 KW Parkinson's disease; dementia; short memory; cancer;
 KW sense or motor function; emotional reaction; fear response; panic;
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 KW tranquiliser.

XX Homo sapiens.
 OS EPI447413-A2.
 FN 18-AUG-2004.
 XX 12-FEB-2004; 2004EP-00003145.
 XX 14-FEB-2003; 2003JP-00102207.
 PR 09-MAY-2003; 2003JP-00131452.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 XX WPI; 2004-583265/57.
 XX P-PSDB; ADR10059.
 XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX Claim 1; SEQ ID NO 1609; 2686pp; English.
 XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cytostatic and tranquiliser activities. This polynucleotide is a full
 CC length human cDNA sequence of the invention. NOTE: This sequence is not
 CC given in the sequence listing of the specification but can be obtained on
 CC CD-ROM from the European Patent Office, Vienna Sub-office.
 XX Sequence 3129 BP; 1002 A; 563 C; 694 G; 870 T; 0 U; 0 Other;
 SQ
 Query Match 17.1%; Score 874.8; DB 13; Length 3129;
 Best Local Similarity 97.1%; Pred No. 4.1e-231;
 Matches 891; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 94 CGAGAACCTTCTGTATACTTAAGAGAGATTTCCAAAATGTGGCCAGATTGCGAGGAGTA 153
 Db 58 CGAGAACCTTCTGTATACTTAAGAGAGATTTCCAAAATGTGGCCAGATTGCGAGGAGTA 117
 QY 154 TCAATATGAGAAAGCTAGGCCATCTGAATACTTTTACTAAGCTTCTTTGTGATATTGC 213
 Db 118 TCAATATGAGAAAGCTAGGCCATCTGAATACTTTTACTAAGCTTCTTTGTGATATTGC 177
 QY 214 CACAGTGAAGAAAACTGGGCTTTTCACTATGAGGATATCATAAATTTGTTGGGTTAGCT 273
 Db 178 CACAGTGAAGAAAACTGGGCTTTTCACTATGAGGATATCATAAATTTGTTGGGTTAGCT 237
 QY 274 TTATTAATGAAGCAAAAGAGTGGCCAGCAGAGGCTTACGAGCGCTTCGATATCTCATC 333
 Db 238 TTATTAATGAAGCAAAAGAGTGGCCAGCAGAGGCTTACGAGCGCTTCGATATCTCATC 297
 QY 334 CAGAGCTCCAGTATTTCTCCAGAAAGGTGCTAAATTTGAAAAGTGGAGCTATTAATAGTAGC 393
 Db 298 CAGAGCTCCAGTATTTCTCCAGAAAGGTGCTAAATTTGAAAAGTGGAGCTATTAATAGTAGC 357
 QY 394 TGCATTGACATACAAACAGAGCAACAGGTTAGAGAGGACACAAGACCTTCGATTAGTCAGA 453
 Db 358 TGCATTGACATACAAACAGAGCAACAGGTTAGAGAGGACACAAGACCTTCGATTAGTCAGA 417

```
QY 454 AAGATGATTACTGGAATGCTTCTCTGTTTCTAGTCTGTGACCAACTATTAAATGCA 513
DB 418 AAGATGATTACTGGAATGCTTCTCTGTTTCTAGTCTGTGACCAACTATTAAATGCA 477
QY 514 GTTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 573
DB 478 GTTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 537
QY 574 TGTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 633
DB 538 TGTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 597
QY 634 TTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 693
DB 598 TTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 657
QY 694 TTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 753
DB 658 TTGAATGATGACATTCAGAGAGACAGAGATGCTCCAGCATGATGCTCCATATC 717
QY 754 AGAATTTAGCACCCTATCTACTGTTTTCATACAGACATAGTCCAGATACAGCTGAAGGA 813
DB 718 AGAATTTAGCACCCTATCTACTGTTTTCATACAGACATAGTCCAGATACAGCTGAAGGA 777
QY 814 CAGCTCAAGAGAGACAGAGAGACAGATTTCTAGCAGTAAATGGAATCATAGCAACA 873
DB 778 CAGCTCAAGAGAGACAGAGAGACAGATTTCTAGCAGTAAATGGAATCATAGCAACA 837
QY 874 TTCCGATCATGGGAGGATTAATTAATTTATGTAACCTGGAATTTCTGGATCCAGTCT 933
DB 838 TTCCGATCATGGGAGGATTAATTAATTTATGTAACCTGGAATTTCTGGATCCAGTCT 897
QY 934 CTAATAGGAGTACTTTGCATACCAAAATATGGAATAGGCGAGTCTACTTGAAGTCTT 993
DB 898 CTAATAGGAGTACTTTGCATACCAAAATATGGAATAGGCGAGTCTACTTGAAGTCTT 957
QY 994 TATGATATTTTCGTCTT 1011
DB 958 TGTTCATCTCGACCT 975

RESULT 11
AAS65246
ID AAS65246 standard; cDNA; 1292 BP.
AC AAS65246;
XX
XX
XX 13-FEB-2002 (first entry)
DE
DE DNA encoding novel human diagnostic protein #1050.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG01059.
XX
```

```
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 1050; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1292 BP; 365 A; 261 C; 312 G; 354 T; 0 U; 0 Other;
Query Match 14.5%; Score 744; DB 5; Length 1292;
Best Local Similarity 88.4%; Pred. No. 4.8e-195;
Matches 879; Conservative 0; Mismatches 0; Indels 115; Gaps 2;
QY 4135 AGGTTTCAAGAGCCTTAAGTTATGATGATTAAGTAAAGAGATTTATTGAGTCTATT 4194
DB 413 AGGTTTCAAGAGCCTTAAGTTATGATGATTAAGTAAAGAGATTTATTGAGTCTATT 472
QY 4195 AATCAAAATACCTGCAACGATCTTCTCAGTGGGTCCATGTTGTCAGTCCACATAT 4254
DB 473 AATCAAAATACCTGCAACGATCTTCTCAGTGGGTCCATGTTGTCAGTCCACATAT 532
QY 4255 GGGGTTTCAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4314
DB 533 GGGGTTTCAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 592
QY 4315 GTAAAGGATATTCCTATTTTCAGACAAATAATACCAACCATGATGATGATGATGAT 4374
DB 593 GTAAAGGATATTCCTATTTTCAGACAAATAATACCAACCATGATGATGATGATGAT 652
QY 4375 AGAGCATTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4434
DB 653 AGAGCATTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
QY 4435 TCTTTTCACTTGTACGACAGCAGATGATGATGATGATGATGATGATGATGATGAT 4494
DB 676 ----- 675
QY 4495 GATGCTCTCTGTTTGTAGAAAGTACAGAGACACTGGACTACAGGAAACATACAGATGAT 4554
DB 676 -----AGAAAGTACAGAGACACTGGACTACAGGAAACATACAGATGAT 718
QY 4555 AACTGCTTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4614
DB 719 AACTGCTTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 778
QY 4615 GCAATATGT-AGTCATTTCAGACTTTCAAGATATTTCCATATTTCTGATTTGGTGAGCAGAC 4673
DB 779 GCAATATGTAGTCAATTCAGACTTTCAAGATATTTCCATATTTCTGATTTGGTGAGCAGAC 838
QY 4674 TATCCATAATTCCTTTAGAAAGTGGTTCCTCTAAAGTTTTCGGGATTTCTGGATGCACTGA 4733
DB -----
```

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 1050; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1292 BP; 365 A; 261 C; 312 G; 354 T; 0 U; 0 Other;

Query Match 14.5%; Score 744; DB 5; Length 1292;
Best Local Similarity 88.4%; Pred. No. 4.8e-195;
Matches 879; Conservative 0; Mismatches 0; Indels 115; Gaps 2;

QY 4135 AGGTTTCAAGAGCCTTAAGTTATGATGATTAAGTAAAGAGATTTATTGAGTCTATT 4194
DB 413 AGGTTTCAAGAGCCTTAAGTTATGATGATTAAGTAAAGAGATTTATTGAGTCTATT 472
QY 4195 AATCAAAATACCTGCAACGATCTTCTCAGTGGGTCCATGTTGTCAGTCCACATAT 4254
DB 473 AATCAAAATACCTGCAACGATCTTCTCAGTGGGTCCATGTTGTCAGTCCACATAT 532
QY 4255 GGGGTTTCAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4314
DB 533 GGGGTTTCAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 592
QY 4315 GTAAAGGATATTCCTATTTTCAGACAAATAATACCAACCATGATGATGATGATGAT 4374
DB 593 GTAAAGGATATTCCTATTTTCAGACAAATAATACCAACCATGATGATGATGATGAT 652
QY 4375 AGAGCATTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4434
DB 653 AGAGCATTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
QY 4435 TCTTTTCACTTGTACGACAGCAGATGATGATGATGATGATGATGATGATGATGAT 4494
DB 676 ----- 675
QY 4495 GATGCTCTCTGTTTGTAGAAAGTACAGAGACACTGGACTACAGGAAACATACAGATGAT 4554
DB 676 -----AGAAAGTACAGAGACACTGGACTACAGGAAACATACAGATGAT 718
QY 4555 AACTGCTTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4614
DB 719 AACTGCTTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 778
QY 4615 GCAATATGT-AGTCATTTCAGACTTTCAAGATATTTCCATATTTCTGATTTGGTGAGCAGAC 4673
DB 779 GCAATATGTAGTCAATTCAGACTTTCAAGATATTTCCATATTTCTGATTTGGTGAGCAGAC 838
QY 4674 TATCCATAATTCCTTTAGAAAGTGGTTCCTCTAAAGTTTTCGGGATTTCTGGATGCACTGA 4733
DB -----

839 TATCCATAATCTTTAGAAAGTGTTCCTCTAAGTTTTTCGGGGAATTTCTCGAATGCAATGA 898
4734 TGGGTGTCTCAAGAGGCTCAGCTAGCAGACCAAAAGCACAGAAATTTGTTACTAGGTGT 4793
899 TGGGTGTCTCAAGAGGCTCAGCTAGCAGACCAAAAGCACAGAAATTTGTTACTAGGTGT 958
4794 TAAACAATTCAGATGATACACCAATGTGCCGTATATCTCTTCGCAAGAAGTCTTAAG 4853
959 TAAACAATTCAGATGATACACCAATGTGCCGTATATCTCTTCGCAAGAAGTCTTAAG 1018
4854 ATTAGTCATTAATTTGAGTAGTTCAGTTCAACTAAATGTGATGAGCTGGGCTTTTAAAC 4913
1019 ATTAGTCATTAATTTGAGTAGTTCAGTTCAACTAAATGTGATGAGCTGGGCTTTTAAAC 1078
4914 AATTAAGGAGAGATATCTCTCAACATTTGATGACATATGCTTTTACTCTGAGGTTTCCCA 4973
1079 AATTAAGGAGAGATATCTCTCAACATTTGATGACATATGCTTTTACTCTGAGGTTTCCCA 1138
4974 TTTGCTGTCACTGCACATTCAGACTTCGGTGTGCGAGGTTTCATACAGAAATTTTCA 5033
1139 TTTGCTGTCACTGCACATTCAGACTTCGGTGTGCGAGGTTTCATACAGAAATTTTCA 1198
5034 AGATGTACAGTTTCTCAAAATGATGCAAGAGCAGAGGCTGTGTTGGCAACACCAACCAA 5093
1199 AGATGTACAGTTTCTCAAAATGATGCAAGAGCAGAGGCTGTGTTGGCAACACCAACCAA 1258
5094 GCAACCTATATGTTGATACATCTCTGCTGAATCCTGA 5127
1259 GCAACCTATATGTTGATACATCTCTGCTGAATCCTGA 1292

RESULT 12
AA196373 standard; cdNA; 765 BP.
XX AC AA196373;
XX DT 13-NOV-2001 (first entry)
XX DE Human neuroblastoma expressed polynucleotide SEQ ID NO 2448.
XX KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX OS Homo sapiens.
XX PN WO200166719-A1.
XX PD 13-SEP-2001.
XX PF 02-MAR-2001; 2001WO-JP001629.
XX PR 07-MAR-2000; 2000JP-00159195.
XX PA (CHIB-) CHIBA PREFECTURE.
XX PA (HISM) HISAMITSU PHARM CO LTD.
XX PI Nakagawara A;
XX DR WPI; 2001-565584/63.
XX PT Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human neuroblastoma,
PT malignancy and susceptibility indicator or tumor marker for anti-cancer
PT agents.
XX PS Claim 1; Page 1801; 2979pp; Japanese.
XX CC The invention relates to novel genes (AA193926-AA197963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes

XX SQ Sequence 765 BP; 231 A; 150 C; 174 G; 200 T; 0 U; 10 Other;
Query Match 12.0%; Score 612.8; DB 4; Length 765;
Best Local Similarity 96.7%; Pred. No. 8.9e-159;
Matches 653; Conservative 0; Mismatches 19; Indels 3; Gaps 3;
QY 94 CGAAGACCTTCTGATAACTTAAAGAGAGATTCTCCAAAATGTGCCAGAGTTGCGAGGATGA 153
DB 71 CGAAGACCTTCTGATAACTTAAAGAGAGATTCTCCAAAATGTGCCAGAGTTGCGAGGATGA 130
QY 154 TCAATATGAGAAAGCTAGGCCATCTGAATAACTTTTAAAGCTTTCTTGTGATATGGC 213
DB 131 TCAATATGAGAAAGCTAGGCCATCTGAATAACTTTTAAAGCTTTCTTGTGATATGGC 190
QY 214 CACAGTGAAGAAAACTGGGCTTTCACTATGAGGATATCATTAATTTGTTTGGGTTAGCT 273
DB 191 CACAGTGAAGAAAACTGGGCTTTCACTATGAGGATATCATTAATTTGTTTGGGTTAGCT 250
QY 274 TTATTAATGAAGCAAAAGAGTCCGAGCAGCAGGGCTACGAGCGCTTCGATATCTCATC 333
DB 251 TTATTAATGAAGCAAAAGAGTCCGAGCAGCAGGGCTACGAGCGCTTCGATATCTCATC 310
QY 334 CAGACTCCAGTATTTCTCCAGAAAGGTGCTTAAATTTGAAAGTGAGCTATTTAAATAGCTAGG 393
DB 311 CAGACTCCAGTATTTCTCCAGAAAGGTGCTTAAATTTGAAAGTGAGCTATTTAAATAGCTAGG 370
QY 394 TGCATTGCACATACAAACAGAGCAACGAGGTAGAGAGGACACAAAGCATTTCGATTAGTCAGA 453
DB 371 TGCATTGCACATACAAACAGAGCAACGAGGTAGAGAGGACACAAAGCATTTCGATTAGTCAGA 430
QY 454 AAGATGATTACTGTGAATGCTTCCTGTTTCTTAGTTCTGTGACCAACTCATTAATTGCA 513
DB 431 AAGATGATTACTGTGAATGCTTCCTGTTTCTTAGTTCTGTGACCAACTCATTAATTGCA 489
QY 514 GTTGGAAATGATGGACTTCAAGAAAGAGACAGAAATGTCGAGCATGCAATTGCCATTATC 573
DB 490 GTTGGAAATGATGGACTTCAAGAAAGAGACAGAAATGTCGAGCATGCAATTGCCATTATC 549
QY 574 TGTGAATCTAGCACTTCAGAAATCCAGAGGTGTTGGCCCTTCGAGGAGGACTAAACACCATA 633
DB 550 TGTGAATCTAGCACTTCAGAAATCCAGAGGTGTTGGCCCTTCGAGGAGGACTAAACACCATA 609
QY 634 TTGAAAAATGTGAT-TGATTTGCCAATTAAGTCGAATAAATGAGGCCCTTAATTACTACAAT 692
DB 610 TTGAAAAATGTGATCCGATTCGCAATTAAGTCGAATAAATGAGGCCCTTAATTACTACAAT 669
QY 693 TTTGCACTCTTCTTAATCATCAAAAGACTCCGAGCATGTGTCGAGCTGATGTAGAATTAGA 752
DB 670 TTTGCACTCTTCTTAATCATCAAAAGACTCCGAGCATGTGTCGAGCTGATGTAGAATTAGA 728
QY 753 GAGAAATTTTAGCACC 767
DB 729 GAGAAATTTTAGCACC 743
RESULT 13
AA196373 standard; cdNA; 1185 BP.
XX AC AA196373;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #1049.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX

PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX P-PSDB; ABG01058.
DR
DR
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX
PS Claim 1; SEQ ID NO 1049; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1185 BP; 360 A; 262 C; 270 G; 293 T; 0 U; 0 Other;

Query Match 10.5%; Score 537.8; DB 5; Length 1185;
Best Local Similarity 85.1%; Pred. No. 7.2e-138;
Matches 663; Conservative 0; Mismatches 2; Indels 114; Gaps 1;

QY 4136 GGTTTCATGAAAGCCTTAAGTTATGTCATCATTTAGATAAAGAGATTATTGAGTCTCTATTA 4195
DB 302 GGTTTCATGAAAGCCTTAAGTTATGTCATCATTTAGATAAAGAGATTATTGAGTCTCTATTA 361

QY 4196 ATCAAAATACCTGCAAGCATCTTCCTCAGTGGCGTCCATGGTCCAGTCCCATATG 4255
DB 362 ATCAAAATACCTGCAAGCATCTTCCTCAGTGGCGTCCATGGTCCAGTCCCATATG 421

QY 4256 GGGGTTTCAGATGATTACATGGTCTCTGCTCCCGGTGGATATAAATGATATATTCAGG 4315
DB 422 GGGGTTTCAGATGATTACATGGTCTCTGCTCCCGGTGGATATAAATGATATATTCAGG 481

QY 4316 TAAAGGATATTCCTATTTTTCAGACAAACCAACATCCACCATGATGATCGAGGTGCAA 4375
DB 482 TAAAGGATATTCCTATTTTTCAGACAAACCAACATCCACCATGATGATCGAGGTGCAA 541

QY 4376 GAGCATTTGCCATGATGTCAGGAGGTCTTCCATCTGGAACTGGAGGTCTTGTAATAAATT 4435
DB 542 GAGCATTTGCCATGATGTCAGG----- 563

QY 4436 CTTTTCACCTGCTACGACAGCATGAGTCTTACGGAATAATGAATTCATCCATTCAG 4495
DB 564 ----- 563

QY 4496 ATGCCTCTCTGTTTTTTAGAAAGTACAGAAGACACTGGAGTACAGAGAACATACAGATGATA 4555
DB 564 -----AGAAAGTACAGAAGACACTGGAGTACAGAGAACATACAGATGATA 607

QY 4556 ACTGCCCTTTATTTGCTCTGTTATGAAATTTCTGGSTTTCCAGCCCGCCAGCAACCAACTGAGTG 4615
DB 608 ACTGCCCTTTATTTGCTCTGTTATGAAATTTCTGGSTTTCCAGCCCGCCAGCAACCAACTGAGTG 667

QY 4616 CAATATGATGATTCATTTCAGACTTTTCAAGATATTCAGATTTCTGATTTGTTGAGCAGACTA 4675
DB 668 CGATATGATGATTCATTTCAGACTTTTCAAGATATTCAGATTTCTGATTTGTTGAGCAGACTA 727

QY 4676 TCCATAATCCTTTAGAAAGTGGTCCCTCTAAGTTTTCGGGGATTTCTGGATGTCAGTATG 4735
DB 728 TCCATAATCCTTTAGAAAGTGGTCCCTCTAAGTTTTCGGGGATTTCTGGATGTCAGTATG 787

QY 4736 GGGTGTCTCAAGAGGCTCAGCTAGCAGCACCACCAAAAGCAGAGATTTCTTACTAGGTCTTA 4795
DB 788 GGGTGTCTCAAGAGGCTCAGCTAGCAGCACCACCAAAAGCAGAGATTTCTTACTAGGTCTTA 847

QY 4796 AAACAAATTCAGATGATACACCAATGTCCTGCTATCTCTTCCGCAAAAGATTTCTTAAGAT 4855
DB 848 AAACAAATTCAGATGATACACCAATGTCCTGCTATCTCTTCCGCAAAAGATTTCTTAAGAT 907

QY 4856 TAGTCATTAATTTGAGTAGTTCAGTTTCAACTAAATGTCATGAGACTGGGCTTTTAAACA 4914
DB 908 TAGTCATTAATTTGAGTAGTTCAGTTTCAACTAAATGTCATGAGACTGGGCTTTTAAACA 966

RESULT 14
AAS80497
ID AAS80497 standard; cDNA; 1185 BP.
XX
AC AAS80497;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16301.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; ABG16310.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 16301; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal

activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1185 BP; 360 A; 262 C; 270 G; 293 T; 0 U; 0 Other;

Query Match 10.5%; Score 537.8; DB 5; Length 1185;
Best Local Similarity 85.1%; Pred. No. 7.2e-138;
Matches 663; Conservative 0; Mismatches 2; Indels 114; Gaps 1;

4136 GGTTCATGAAGCCTTAAGTTATGATCATTAAGATAAAGAGATTTATGATCCTATTA 4195
|||||
302 GGTTCATGAAGCCTTAAGTTATGATCATTAAGATAAAGAGATTTATGATCCTATTA 361
|||||
4196 ATCAAAATACCCCTGCAACGATCTTCCTCAGTGGTCCATGTTCCAGTGCCACATATG 4255
|||||
362 ATCAAAATACCCCTGCAACGATCTTCCTCAGTGGTCCATGTTCCAGTGCCACATATG 421
|||||
4256 GGGGTTCCAGATGATTAATGCTTGTCTCCCGTGGATTAATGATATATATCCAGG 4315
|||||
422 GGGGTTCCAGATGATTAATGCTTGTCTCCCGTGGATTAATGATATATATCCAGG 481
|||||
4316 TAAAGGATATTCCTATTTTCAGACAAAACATACACCATGATGATGAGGTGCAA 4375
|||||
482 TAAAGGATATTCCTATTTTCAGACAAAACATACACCATGATGATGAGGTGCAA 541
|||||
4376 GAGCATTTGCCATGATGAGGAGGTCTCCATCTGGAAGTGGAGTCTTTGAAAAATTT 4435
|||||
542 GAGCATTTGCCATGATGAGGAGGTCTCCATCTGGAAGTGGAGTCTTTGAAAAATTT 563
|||||
4436 CTTTTTCATCTGCTACGACAGCAGATGATCTTACGGAAATATGAATTCATCCATCAG 4495
|||||
564 ----- 563

4496 ATGCTCTCTGTTTATGAAGTACAGACACTGACATACAGGAAACATACAGATGATA 4555
|||||
564 -----AGAAAGTACAGAGACACTGACATACAGGAAACATACAGATGATA 607
|||||
4556 ACTGCTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4615
|||||
608 ACTGCTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 667
|||||
4616 CAATATGATGATTCATGACACTTTCAGATATTCATATTCCTGATTTGGTGAGCAGACTA 4675
|||||
668 CGATATGATGATTCATGACACTTTCAGATATTCATATTCCTGATTTGGTGAGCAGACTA 727
|||||
4676 TCCATAATCCCTTTAGAGTGGTTCCTCTAAGTTTTCGGGATTTCTGGATGAGTATG 4735
|||||
728 TCCATAATCCCTTTAGAGTGGTTCCTCTAAGTTTTCGGGATTTCTGGATGAGTATG 787
|||||
4736 GGGTGTCTCAGAGGCTCAGCTAGCAGCAGCAAAAGCAGAGATTTGTTAGTGGTGTTA 4795
|||||
788 GGGTGTCTCAGAGGCTCAGCTAGCAGCAGCAAAAGCAGAGATTTGTTAGTGGTGTTA 847
|||||
4796 AAACAATTTCCAGATGATACCAATGTCCTATCTCTTTCGCAAGAGATTTCTAAGAT 4855
|||||
848 AAACAATTTCCAGATGATACCAATGTCCTATCTCTTTCGCAAGAGATTTCTAAGAT 907
|||||
4856 TAGTCATTAATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGATTTTACA 4914
|||||
908 TAGTCATTAATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGATTTTACA 966
|||||

RESULT 15
ADC32002
ID ADC32002 standard; cDNA; 1185 BP.
XX
AC ADC32002;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA contig sequence, SEQ ID NO:2084.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 5; ss.
XX
OS Homo sapiens.
XX
FN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
XX
PT P-PSDB; ADC32769.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX
PS Example 2; SEQ ID NO 2084; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides

CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig
CC sequence used in an example of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 1185 BP; 360 A; 262 C; 270 G; 293 T; 0 U; 0 Other;

```
Query Match      10.5%; Score 537.8; DB 10; Length 1185;
Best Local Similarity 85.1%; Pred. No. 7.2e-138;
Matches 663; Conservative 0; Mismatches 2; Indels 114; Gaps 1;

Qy 4136 GGTTCATGAAGCCTTAAGTTATGATCATTCATTAGATAAAGAGATTTATTGAGTCTCTATTA 4195
Db      |||||
Qy 302  GGTTCATGAAGCCTTAAGTTATGATCATTCATTAGATAAAGAGATTTATTGAGTCTCTATTA 361
Db      |||||
Qy 4196 ATCAAAATACCCGCAAGCATCTTCCTCAGTGGGTCCAGTCCAGTCCACATATG 4255
Db      |||||
Qy 362  ATCAAAATACCCGCAAGCATCTTCCTCAGTGGGTCCAGTCCAGTCCACATATG 421
Db      |||||
Qy 4256 GGGGTTTCAGATGATTACATTTGCTCTCTCCCGTGGATATAAATGATATATTTCCAGG 4315
Db      |||||
Qy 422  GGGGTTTCAGATGATTACATTTGCTCTCTCCCGTGGATATAAATGATATATTTCCAGG 481
Db      |||||
Qy 4316 TAAAGGATATTTCCCTATTTTCAGACAAATAACATACCAACATCATGATCGAGGTGCAA 4375
Db      |||||
Qy 482  TAAAGGATATTTCCCTATTTTCAGACAAATAACATACCAACATCATGATCGAGGTGCAA 541
Db      |||||
Qy 4376 GAGCATTTGCCCATGATGAGGAGGTCTTCATCTGGAACTGGAGGCTCTGTAAAAAATT 4435
Db      |||||
Qy 542  GAGCATTTGCCCATGATGAGG----- 563
Qy 4436 CTTTTCACCTGCTACGACAGCATGAGTCTTACGGAAATAATGAATTCATCCATTCAG 4495
Db      |||||
Qy 564  ----- 563
Qy 4496 ATGCCTCTCTGTTTTTAGAAGTACAGAAGACACTGGACTACAGGAACATACAGATGATA 4555
Db      |||||
Qy 564  -----AGAAAGTACAGAAGACACTGGACTACAGGAACATACAGATGATA 607
Qy 4556 ACTGCCCTTATTTGTCGTGATTTGAAATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTG 4615
Db      |||||
Qy 608  ACTGCCCTTATTTGTCGTGATTTGACATTTCTGGGTTTCCAGCCAGCAACCAACTGAGTG 667
Qy 4616 CAATATGTAGTCAATTCAGACTTTTCAAGATATTTCCATATTTCTGATTTGTTGAGCAGACTA 4675
Db      |||||
Qy 668  CGATATGTAGTCAATTCAGACTTTTCAAGATATTTCCATATTTCTGATTTGTTGAGCAGACTA 727
Qy 4676 TCCATAATCCCTTTAGAAAGTGGTTCCTCTAAGTTTTCGGGGATTTCTGGATGCACTGATG 4735
Db      |||||
Qy 728  TCCATAATCCCTTTAGAAAGTGGTTCCTCTAAGTTTTCGGGGATTTCTGGATGCACTGATG 787
Qy 4736 GGGTGTCTCAAGAAGGCTCAGCTAGCAGACCAAAAGCAGCAAAATTTGTTACTAGGTGTTA 4795
Db      |||||
Qy 788  GGGTGTCTCAAGAAGGCTCAGCTAGCAGACCAAAAGCAGCAAAATTTGTTACTAGGTGTTA 847
Qy 4796 AAACAATTCAGATGATACACCAATGTCCTTACTCTCCGCAAGAGATTTCTAAGAT 4855
Db      |||||
Qy 848  AAACAATTCAGATGATACACCAATGTCCTTACTCTCCGCAAGAGATTTCTAAGAT 907
Qy 4856 TAGTCATTAATTTGAGTAGTTGAGTTTCAACTAAATGTCATGAGACTGGGCTTTTAAACA 4914
Db      |||||
Qy 908  TAGTCATTAATTTGAGTAGTTGAGTTTCAACTAAATGTCATGAGACTGGGCTTTTAAACA 966
```

Search completed: April 8, 2005, 19:02:16
Job time : 1718.98 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 17:20:47 ; Search time 550.17 Seconds
(without alignments)
15248.348 Million cell updates/sec

Title: US-10-782-244-2

Perfect score: 5127

Sequence: 1 atggcgcgatcgccgagg.....atacatctgtgaatcctga 5127

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370	7.2	401	US-09-513-999C-942	Sequence 942, App
2	62.4	1.2	447	US-09-270-767-1698	Sequence 1698, App
3	62.4	1.2	447	US-09-270-767-16980	Sequence 16980, A
C 4	60	1.2	1141	US-09-806-708B-22	Sequence 22, Appl
5	58.6	1.1	1141	US-09-806-708B-22	Sequence 22, Appl
C 6	53	1.0	7218	US-08-232-463-14	Sequence 14, Appl
7	44	0.9	399	US-09-621-976-8976	Sequence 8976, App
C 8	41	0.8	120727	US-09-949-016-15787	Sequence 15787, A
C 9	41	0.8	120727	US-09-949-016-15788	Sequence 15788, A
C 10	40.8	0.8	832	US-09-621-976-2813	Sequence 2813, App
C 11	39.2	0.8	252	US-09-134-000C-532	Sequence 532, App
12	39.2	0.8	3247	US-09-176-664-19	Sequence 19, Appl
13	39.2	0.8	3319	US-09-176-664-17	Sequence 17, Appl
C 14	39	0.8	83665	US-09-949-016-16995	Sequence 16995, A
C 15	38.8	0.8	2051	US-09-636-215-699	Sequence 699, App
16	38.8	0.8	2051	US-09-685-166A-699	Sequence 699, App
17	38.8	0.8	2051	US-09-679-426-699	Sequence 699, App
18	38.8	0.8	2051	US-09-759-143-899	Sequence 699, App
19	38.8	0.8	2051	US-09-651-236-699	Sequence 699, App
20	38.8	0.8	10478	US-09-445-774-16	Sequence 16, Appl
C 21	38.8	0.8	256171	US-09-949-016-12822	Sequence 12822, A
C 22	38.8	0.8	256176	US-09-949-016-15524	Sequence 15524, A
C 23	38	0.7	861	US-09-976-594-289	Sequence 289, App
C 24	38	0.7	1664976	US-08-916-421B-1	Sequence 1, Appli
C 25	38	0.7	1664976	US-09-692-570-1	Sequence 1, Appli
26	37.8	0.7	505	US-09-621-976-15639	Sequence 15639, A
27	37.8	0.7	1860	US-09-919-497-50	Sequence 50, Appl

C 28	37.8	0.7	52494	4	US-09-949-016-16498	Sequence 16498, A
C 29	37.8	0.7	94095	4	US-09-949-016-14389	Sequence 14389, A
C 30	37.8	0.7	278866	4	US-09-949-016-13922	Sequence 13922, A
31	37.8	0.7	278866	4	US-09-949-016-13923	Sequence 13923, A
32	37.8	0.7	278866	4	US-09-949-016-13924	Sequence 13924, A
33	37.8	0.7	278866	4	US-09-949-016-13925	Sequence 13925, A
34	37.8	0.7	278866	4	US-09-949-016-13926	Sequence 13926, A
35	37.8	0.7	278866	4	US-09-949-016-14699	Sequence 14699, A
36	37.8	0.7	278866	4	US-09-949-016-14700	Sequence 14700, A
37	37.8	0.7	278866	4	US-09-949-016-14701	Sequence 14701, A
38	37.8	0.7	278866	4	US-09-949-016-14702	Sequence 14702, A
39	37.8	0.7	278866	4	US-09-949-016-14703	Sequence 14703, A
40	37.6	0.7	10274	4	US-09-827-688-5	Sequence 5, Appli
C 41	37.6	0.7	90724	4	US-09-949-016-16601	Sequence 16601, A
C 42	37.4	0.7	601	4	US-09-949-016-179262	Sequence 179262, A
C 43	37.4	0.7	263693	4	US-09-949-016-12386	Sequence 12386, A
C 44	37.4	0.7	263694	4	US-09-949-016-16915	Sequence 16915, A
45	37.2	0.7	2069	4	US-09-949-016-5801	Sequence 5801, Ap

ALIGNMENTS

RESULT 1

US-09-513-999C-942
; Sequence 942, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 942
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..400
US-09-513-999C-942

Query Match 7.2%; Score 370; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 7.3e-102; Mismatches 0; Gaps 0;
Matches 370; Conservative 0; Indels 0; Gaps 0;

Qy	1	ATGCGCGCGATCGCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGCGCGGAATGAC	60
Db	32	ATGCGCGCGATCGCGCGCGCGCTCTCTGAAGAACCTCCGAGTACGAGCGCGGAATGAC	91
Qy	61	AGCGCGGAGGAGAACGTCCTCGGATCTGACCGGAGAACCTTCTGATACTTAAGAGAG	120
Db	92	AGCGCGGAGGAGAACGTCCTCGGATCTGACCGGAGAACCTTCTGATACTTAAGAGAG	151
Qy	121	ATTCTCCAAAATGTGGCCAGATTCGAGGAGATTCAAATATGAGAAAGCTAGGCCATCTG	180
Db	152	ATTCTCCAAAATGTGGCCAGATTCGAGGAGATTCAAATATGAGAAAGCTAGGCCATCTG	211
Qy	181	AATAACTTTTACTAAGCTTCTTTGTGATTTGGCCACAGTGAAGAAAACCTTTGGCTTTTAC	240
Db	212	AATAACTTTTACTAAGCTTCTTTGTGATTTGGCCACAGTGAAGAAAACCTTTGGCTTTTAC	271
Qy	241	TATGAGATATCATTAATTTCTTTCGGTTAGCTTTTATTAATGAAGAAAAGAGTGGGA	300
Db	272	TATGAGATATCATTAATTTCTTTCGGTTAGCTTTTATTAATGAAGAAAAGAGTGGGA	331

[illegible]

```

RESULT 6
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:

```


	Best Local Similarity	49.3%;	Pred: No. 10;	Mismatches	0;	Mismatches	110;	Indels	0;	Gaps	0;
	Matches	107;	Conservative								
Qy	958	AATATGAAATAAGCGAGGTCTACTTGAAGTGCTTTATGATATATTTCGTCTTCCCTCTA	1017								
Db	28912	AAGCTGCACCTTTGGGAGGCCACCCCAGAGACGAGAAGTTTAGCTTCTCCTAGCTGA	28853								
Qy	1018	CCTGTGTTGTCACTCAGGAGTTTCATAGAAGCACACTACTCAGTGTAGATCCAGGGAGGTTCCAA	1077								
Db	28852	GGGGTGGGAGGGTGCAGGGGTTCCACGCTCAGGGAGGAGAGGAGGGGTGATNCT	28793								
Qy	1078	GACAGTGGAGGCTTTTCAGATGGCTTTTGGCAGCTGAGGCAAAAACATATTCTTCCCTCAT	1137								
Db	28792	TTCACCTGAGGTTTTCTTCTGAGGCTGTGGCTGTGGAGGTCCTACTCCTCTCCTGAG	28733								
Qy	1138	CGTGCCAGATCCAGGCCAGACCTCATGGATAATTATT	1174								
Db	28732	CCCTCTCTACCACAGACTGGCTCTCCAGCGAGTTCTT	28696								

```

RESULT 9
US-09-949-016-15788/c
; Sequence 15788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15788
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15788

```

	Query Match	0.8%	Score 41;	DB 4;	Length 120727;
	Best Local Similarity	49.3%	Pred. No. 10;		
	Matches 107;	Conservative 0;	Mismatches 110;	Indels 0;	Gaps 0;
QY	958	AATATCGAAATAGGCGAGGTCTACTTTGAAGTGTCTTTATGATATATTTTCGTCTTCCTCTA	1017		
DB	28912	AGCTGGCACCTTTCGAGGGCCNCCCAGAGGACAGAGTTTAGCTTCCTCTAGCTGA	28953		
QY	1018	CTGTGTTGTGACTGAGGAGTTTCATAGAAAGCACTACTCAGTGTAGATCCAGGGGAGGTTCCAA	1077		
DB	28852	GGGGTGGGAGGGTGAGGGGTTCCAGCTCAGGGGAGGACAGAGGAAGGGTGTGATTC	28793		
QY	1078	GACAGTTTGAGGGCTTTTCAGATGCTTTTGGCGAGCTGAGGCAAAACTATTCTTCCTCAT	1137		
DB	28792	TTCACTCTGAGGTTTTCTTGAGGCTGTGGCTGTGGAGGGTCTACTCCCTCTTCTGAG	28733		
QY	1138	CGTGCCAGATCCAGGCCAGACCTCTCATGGATAATTAT	1174		
DB	28732	CCCTCTCTACACAGACTGGCTCTCCAGCGAGTCTT	28696		

RESULT 10
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976

```

; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

```

	Query Match	0.8%	Score 40.8	DB 4	Length 832
	Best Local Similarity	9.3%	Pred. No. 0.28		
	Matches	32	Conservative 181	Mismatches 128	Indels 3
			Gaps 3		
Qy	3249	TTCAATCCCTTCCTTCCTTCCTAGTAAACTT-GTGAAGATCGTATCTCTAAATTCGCTT	3306		
Db	360	WTWKTKTWYVTVTTRMMWKKRWRYYWVKSTYACASRYKRYTGWYMYWKKRMNST	301		
Qy	3307	ACTTTCGCTTAAACAAAAACATCTAGTAGCAGTGCCTCAAGAGGAGGAGAAATTTATCATCT	3366		
Db	300	RWTCYMCWCKCYRGRCAYTWARGRMWSYANGKWSMRSANSMCTRYMYKGSYWTM	241		
Qy	3367	GAAGGTGAAGACAGCAACAGGCGGAATCAGAACCTTTACGGAGCCCATGTGTTGATTTAAT	3426		
Db	240	KCTCATWCYMYKYWYKRMWSKTCWSGSGGYMTSYSTRSYWASWMYTMCWNGRW	181		
Qy	3427	CATAGTGATGATTTTACACCCATATCCATCTGTACAGAAACATTACAAATTAGAGACTTCA	3486		
Db	180	STYYWAGCKWRYATTTWRRAMWWAAWTMMYMWAWCWSSSRGAAMYRTTMMWGYR	121		
Qy	3487	TTTATCGGGAATAAGACACATTTGAAGACACTGGTGTAGTACACCAAGCATTCGAGAAAAATGAC	3546		
Db	120	YWR- KKSRYRTCAWAYAWKTBSYYWCWKWKRCMMMMMAWYKGTMMWRACWKTR	62		
Qy	3547	TTAAATTCACCAAGAAATTTTGGTACAGAGAATTCACAGAGAAA	3590		
Db	61	YWRWAWAWRWMTTMMMYYYWTRAKRRWYWRKWRWSWMM	18		

RESULT 11
US-09-134-000C-532
; Sequence 532, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 532
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (96)..(126)
; OTHER INFORMATION: Nucleotides 96, 98, 105, 106 & 126 are "n" wherein
; OTHER INFORMATION: "n" = any nucleotide.

```
; OTHER INFORMATION: 1
US-09-134-000C-532

Query Match      0.8%; Score 39.2; DB 4; Length 252;
Best Local Similarity 51.6%; Pred. No. 0.36;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 130 AATGTGGCCAGATGTCAGGAGATCAATATATGAGAAAGCTAGCCATCTGAATACTTT 189
    |||||
Db 87 AATGGGTCTWANTACCGANNAGTCTAAATCTAGATTTTNAAGTTGATTAAGATGATTT 146

QY 190 ACTAAGCTCTTTGTGATATGGCCACAGTGAAGAAACTGGGCTTTTCATATGAGGAT 249
    |||||
Db 147 TCTGAGGCCCTTTGAAATCCTTCGAAAATTAATGAATTTGGTTCTCACCTTAAGACGAA 206

QY 250 ATCATAAATTTGTTGGGTTAGCTTTTATTAATGAAGCAAA 290
    |||||
Db 207 ARCAATGATGTTTGGCGTTTCAATTTATCAACTTACGGAAA 247

RESULT 12
US-09-176-664-19
; Sequence 19, Application US/09176664
; Patent No. 6815539
; GENERAL INFORMATION:
; APPLICANT: Salkoff, Lawrence
; APPLICANT: Schreiber, Matthew
; APPLICANT: Silvia, Chris
; APPLICANT: The Washington University
; APPLICANT: ICAGEN Inc.
; TITLE OF INVENTION: A pH Sensitive Potassium Channel in Spermatoocytes
; FILE REFERENCE: 018512-000120US
; CURRENT APPLICATION NUMBER: US/09/176,664
; PRIOR FILING DATE: 1998-10-21
; EARLIER APPLICATION NUMBER: US 60/063,138
; EARLIER FILING DATE: 1997-10-22
; EARLIER APPLICATION NUMBER: US 60/076,172
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 3319
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human slo3-1 (hslo3-1)
US-09-176-664-17

Query Match      0.8%; Score 39.2; DB 4; Length 3319;
Best Local Similarity 49.5%; Pred. No. 2.5;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 3759 AAGCATGAGTACTGTGGTAAGTACTAAACTATTAAAGCAAGCCACTATTTCACGCCACA 3818
    |||||
Db 2421 AAGCAACACAGACTTTGGTAGACACAGAGCCATCATGCGAACCCCTCACCATCGATCCTT 2480

QY 3819 GTCTAACCATCTGCTCTCTCCAAATCAAAATTCGGTGTCCTCCAGGTTCTTC 3878
    |||||
Db 2481 GCAAAATTGACTCTCTCTGACCCCTGACCCCTCAGTGTTCAGAGGAGACTCCAGGTTACAC 2540

QY 3879 TCATACGCTTCCTAGAAGACAGCAGTCCCTTAAAGCACCCCTCTATTGCTACAAATTAAG 3938
    |||||
Db 2541 AAATGGACATATGAGAAATCAAACTGCGGAAAAGTCCCTTACTGAACTGAATAA 2600

QY 3939 TCTAGCAGATTGTAACCTTTAGTTA 3962
    |||||
Db 2601 TCCTTCCAACATTCACCTTTATTGA 2624

RESULT 14
US-09-949-016-16995/c
; Sequence 16995, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16995
; LENGTH: 83665
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: human slo3-2 (hslo3-2)
US-09-949-016-16995

Query Match      0.8%; Score 39; DB 4; Length 83665;
Best Local Similarity 46.8%; Pred. No. 31;
Matches 123; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
```

```
; OTHER INFORMATION: 1
US-09-134-000C-532

Query Match      0.8%; Score 39.2; DB 4; Length 252;
Best Local Similarity 51.6%; Pred. No. 0.36;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 130 AATGTGGCCAGATGTCAGGAGATCAATATATGAGAAAGCTAGCCATCTGAATACTTT 189
    |||||
Db 87 AATGGGTCTWANTACCGANNAGTCTAAATCTAGATTTTNAAGTTGATTAAGATGATTT 146

QY 190 ACTAAGCTCTTTGTGATATGGCCACAGTGAAGAAACTGGGCTTTTCATATGAGGAT 249
    |||||
Db 147 TCTGAGGCCCTTTGAAATCCTTCGAAAATTAATGAATTTGGTTCTCACCTTAAGACGAA 206

QY 250 ATCATAAATTTGTTGGGTTAGCTTTTATTAATGAAGCAAA 290
    |||||
Db 207 ARCAATGATGTTTGGCGTTTCAATTTATCAACTTACGGAAA 247

RESULT 12
US-09-176-664-19
; Sequence 19, Application US/09176664
; Patent No. 6815539
; GENERAL INFORMATION:
; APPLICANT: Salkoff, Lawrence
; APPLICANT: Schreiber, Matthew
; APPLICANT: Silvia, Chris
; APPLICANT: The Washington University
; APPLICANT: ICAGEN Inc.
; TITLE OF INVENTION: A pH Sensitive Potassium Channel in Spermatoocytes
; FILE REFERENCE: 018512-000120US
; CURRENT APPLICATION NUMBER: US/09/176,664
; PRIOR FILING DATE: 1998-10-21
; EARLIER APPLICATION NUMBER: US 60/063,138
; EARLIER FILING DATE: 1997-10-22
; EARLIER APPLICATION NUMBER: US 60/076,172
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 3247
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human slo3-2 (hslo3-2)
US-09-176-664-19

Query Match      0.8%; Score 39.2; DB 4; Length 3247;
Best Local Similarity 49.5%; Pred. No. 2.4;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 3759 AAGCATGAGTACTGTGGTAAGTACTAAACTATTAAAGCAAGCCACTATTTCACGCCACA 3818
    |||||
Db 2349 AAGCAACACAGACTTTGGTAGACACAGAGCCATCATGCGAACCCCTCACCATCGATCCTT 2408

QY 3819 GTCTAACCATCTGCTCTCTCCAAATCAAAATTCGGTGTCCTCCAGGTTCTTC 3878
    |||||
Db 2409 GCAAAATTGACTCTCTCTGACCCCTGACCCCTCAGTGTTCAGAGGAGACTCCAGGTTACAC 2468

QY 3879 TCATACGCTTCCTAGAAGACAGCAGTCCCTTAAAGCACCCCTCTATTGCTACAAATTAAG 3938
    |||||
Db 2469 AAATGGACATATGAGAAATCAAACTGCGGAAAAGTCCCTTACTGAACTGAATAA 2528

QY 3939 TCTAGCAGATTGTAACCTTTAGTTA 3962
    |||||
Db 2529 TCCTTCCAACATTCACCTTTATTGA 2552

RESULT 13
US-09-176-664-17
; Sequence 17, Application US/09176664
; Patent No. 6815539
```

QY 4855 TTAGTCATTAATTGAGTAGTTAGTTCACTAAATGTCAGAGACTGGCGCTTTTAAACA 4914
 |||||
 Db 19461 TTTCGATTTTATTGATATGATCATTTTCTTCAATATTGGATGTGATCTGCTTGTGTTT 19402
 |||||
 QY 4915 ATTAAGGAGAGTATCCTCAACATTTGATGACATATGCCCTTACTCTGAGGTTTCCCAT 4974
 |||||
 Db 19401 ATTTTAGATCAATTTCTTCAATATTGAGTCCATATCACATTCTCCTAAAGAGTTTTCATT 19342
 |||||
 QY 4975 TTGCTGTCACTGCACATTCAGACTTCCGTGTGGAGGTTTCAACAAGAAATTTTCAA 5034
 |||||
 Db 19341 ATTAACAAGCTCTCCATTTAATTCATTAGATTTTAAAGCCTTAAAGAGTAACCTTA 19282
 |||||
 QY 5035 GATGTACAGTTTCTACAAATGATGAAGAAGAGGCTGTGTGGCAACACCAACCAAG 5094
 |||||
 Db 19281 GAATTTAAATTAGATGTACTTGTCTGATATACCATCTAATATAAATTCAAACATTTTACATG 19222
 |||||
 QY 5095 CAACCTATAGTTGATACATCTGC 5117
 |||||
 Db 19221 CTACAAATATTGCTAGAAGTAC 19199
 |||||

RESULT 15

US-09-636-215-699
 ; Sequence 699, Application US/09636215
 ; Patent No. 6620922
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqui
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stoik, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.42717C17
 ; CURRENT APPLICATION NUMBER: US/09/636,215
 ; CURRENT FILING DATE: 2000-08-10
 ; NUMBER OF SEQ ID NOS: 852
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 699
 ; LENGTH: 2051
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(2051)
 ; OTHER INFORMATION: n = A,T,C or G

Query Match 0.8%; Score 38.8; DB 4; Length 2051;
 Best Local Similarity 56.2%; Pred. No. 2.3;
 Matches 73; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
 QY 1978 TTTTATTATTGAACAACATTTCTTGGCACCCCTCATGGAGTTAAAAATGCTGGAAAAATGC 2037
 |||||
 Db 1713 TTACTATCAATTGAAGCATATTTTCAATTATTCTGTTAATAATATGGAGGAGGATGA 1772
 |||||
 QY 2038 AGTGTATTTCAGTCTCTCCTTAATCTTTGCTCTTGAACCAAGATCACTTGTCTAAA 2097
 |||||
 Db 1773 AATTGTTTTTATTCTTTTAGAATTTTTTTTTTATCAGGAAAAACAGAGGTAAAGTGTATCA 1832
 |||||

QY 2098 CTTACTGTGTTT 2107
 |||||
 Db 1833 ATTACTATTT 1842
 |||||
 Search completed: April 9, 2005, 13:48:39
 Job time : 555.17 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 17:46:42 ; Search time 1846.73 Seconds
(without alignments)
16824.125 Million cell updates/sec

Title: US-10-782-244-2
Perfect score: 5127
Sequence: 1 atggcgccgcatcgccgcg.....atacatctgtgaatctga 5127

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT NEW PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5127	100.0	5127	18	US-10-782-244-2
2	5127	100.0	5662	18	US-10-782-244-1
3	1772.8	34.6	4657	10	US-09-814-353-19068
C 4	1768	34.5	4601	18	US-10-357-930-20181
C 5	1768	34.5	4602	18	US-10-357-930-26008
6	465.2	9.1	570	10	US-09-814-353-11
7	465.2	9.1	570	10	US-09-814-353-6409
C 8	385.6	7.5	452	18	US-10-357-930-9378
C 9	383.4	7.5	427	18	US-10-357-930-30581
C 10	383.4	7.5	427	18	US-10-357-930-39551
C 11	336.4	6.6	416	18	US-10-357-930-209

12	334	6.5	341	10	US-09-814-353-12794	Sequence 12794, A
C 13	321	6.3	450	9	US-09-933-797-314	Sequence 714, Appl
14	280.2	5.5	726	9	US-09-910-943-73	Sequence 313, Appl
15	166.8	3.3	567	16	US-10-029-386-8811	Sequence 8811, Ap
16	163.6	3.2	170	16	US-10-029-386-22537	Sequence 22537, A
17	142.2	2.8	1698	13	US-10-027-632-257900	Sequence 257900,
18	142.2	2.8	1698	17	US-10-027-632-257900	Sequence 257900,
19	88.6	1.7	405	17	US-10-242-535A-44306	Sequence 44306, A
20	88.6	1.7	405	17	US-10-085-783A-44306	Sequence 44306, A
21	77.4	1.5	289	10	US-09-814-353-18070	Sequence 18070, A
22	73	1.4	185	10	US-09-814-353-21600	Sequence 21600, A
C 23	65.2	1.3	170	17	US-10-242-535A-22309	Sequence 22309, A
C 24	65.2	1.3	170	17	US-10-085-783A-22309	Sequence 22309, A
C 25	60.8	1.2	280	18	US-10-674-124A-8321	Sequence 8321, Ap
26	57	1.1	185	10	US-09-814-353-5399	Sequence 5399, Ap
27	57	1.1	185	10	US-09-814-353-11686	Sequence 11686, A
28	51.2	1.0	530	18	US-10-437-963-76169	Sequence 76169, A
C 29	41.8	0.8	37515	18	US-10-433-793-28	Sequence 28, Appl
C 30	41	0.8	454	10	US-09-918-995-15405	Sequence 15405, A
C 31	41	0.8	1359	13	US-10-027-632-166149	Sequence 166149,
C 32	41	0.8	1359	17	US-10-027-632-166149	Sequence 166149,
C 33	41	0.8	136726	17	US-10-085-117-244	Sequence 244, App
34	40.6	0.8	1136	9	US-09-938-842A-3011	Sequence 3011, Ap
35	40.6	0.8	1136	11	US-09-938-842A-3011	Sequence 3011, Ap
36	40.4	0.8	3576	17	US-10-282-122A-15864	Sequence 15864, A
C 37	40.4	0.8	6384	15	US-10-311-455-1897	Sequence 1897, Ap
C 38	40	0.8	14603	19	US-10-741-600-17799	Sequence 17799, A
C 39	39.8	0.8	533	13	US-10-027-632-221849	Sequence 221849,
40	39.8	0.8	533	17	US-10-027-632-221849	Sequence 221849,
C 41	39.8	0.8	12069	17	US-10-257-166-11	Sequence 11, Appl
42	39.6	0.8	232	18	US-10-437-963-97297	Sequence 97297, A
43	39.4	0.8	655	11	US-09-969-034-2726	Sequence 2726, Ap
44	39.2	0.8	6061	15	US-10-311-455-113	Sequence 113, App
C 45	39.2	0.8	6175	15	US-10-311-455-1280	Sequence 1280, Ap

ALIGNMENTS

RESULT 1
US-10-782-244-2
; Sequence 2, Application US/10782244
; Publication No. US20040253677A1
; GENERAL INFORMATION:
; APPLICANT: Sabatini, David M. ✓
; APPLICANT: Kim, Do-Hyung
; APPLICANT: Sarbasov, Dos D.
; TITLE OF INVENTION: MOTOR KINASE-ASSOCIATED PROTEINS
; FILE REFERENCE: WBL-P01-010
; CURRENT APPLICATION NUMBER: US/10/782,244
; CURRENT FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/448,035
; PRIOR FILING DATE: 2003-02-18
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 5127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-782-244-2

Query Match	100.0%	Score 5127;	DB 18;	Length 5127;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 5127;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGCGCGCATCGCGCGCGCGCTCTCTGAAGAACCTTCGAGTACGAGCGCGAATGAC	60	
Db	1	ATGCGCGCATCGCGCGCGCGCTCTCTGAAGAACCTTCGAGTACGAGCGCGAATGAC	60	
Qy	61	AGCGCGGAGGAGAACCTCCGCTGGATCTGACCCGAGAACCTTCTGTAACCTTAAGAGAG	120	
Db	61	AGCGCGGAGGAGAACCTCCGCTGGATCTGACCCGAGAACCTTCTGTAACCTTAAGAGAG	120	

QY 121 ATTCTCCAAATGTGGCCAGATTGCGGAGATCAAAATATGAGAAAGCTAGGCCATCTG 180
 Db 121 ATTCTCCAAATGTGGCCAGATTGCGGAGATCAAAATATGAGAAAGCTAGGCCATCTG 180
 QY 181 AATAACTTTACTAAGCTCTTTGTGATATGGCCACAGTGAAGAAAGCTGGCTTTCAC 240
 Db 181 AATAACTTTACTAAGCTCTTTGTGATATGGCCACAGTGAAGAAAGCTGGCTTTCAC 240
 QY 241 TATGAGATATCATATTTGTTGGGTAGCTTTTAAATGAAGCAAAAGAGTGCGA 300
 Db 241 TATGAGATATCATATTTGTTGGGTAGCTTTTAAATGAAGCAAAAGAGTGCGA 300
 QY 301 GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGATATCTCCAGAGGTG 360
 Db 301 GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGATATCTCCAGAGGTG 360
 QY 361 CTAAATTTGAAAGTGACTATTTAATAGCTAGTGCATTCAGATATCAACAGAGCAACGAG 420
 Db 361 CTAAATTTGAAAGTGACTATTTAATAGCTAGTGCATTCAGATATCAACAGAGCAACGAG 420
 QY 421 GTAGAGGACACAGCACTTCGATTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG 480
 Db 421 GTAGAGGACACAGCACTTCGATTAGTCAGAAAGATGATTACTGTGAATGCTTCCTTG 480
 QY 481 TTTCTAGTTCTGACCAACTCATTTAATGTCAGTTGGAATGATGGACTTCAAGAAAGA 540
 Db 481 TTTCTAGTTCTGACCAACTCATTTAATGTCAGTTGGAATGATGGACTTCAAGAAAGA 540
 QY 541 GACAGAAATGTCGAGCAGTGCATTCGATTAATCTGTGAATAGCACTTCAGAAATCCAGAG 600
 Db 541 GACAGAAATGTCGAGCAGTGCATTCGATTAATCTGTGAATAGCACTTCAGAAATCCAGAG 600
 QY 601 GTGGTGCCCTTCGAGGAGACTAACAACATATGAAATGATGATGATGATGATGATGATG 660
 Db 601 GTGGTGCCCTTCGAGGAGACTAACAACATATGAAATGATGATGATGATGATGATGATG 660
 QY 661 AGTCGAATAATGAGGCCCTTAATTAATCAATTTGACCTTCTTAATCATCCAAAGACT 720
 Db 661 AGTCGAATAATGAGGCCCTTAATTAATCAATTTGACCTTCTTAATCATCCAAAGACT 720
 QY 721 CGGCAGTATGTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 Db 721 CGGCAGTATGTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 QY 781 CACTACAGATAGTCAGATACAGCTGAAGGACAGCTCAAGAGACAGAGAGAGCAGCA 840
 Db 781 CACTACAGATAGTCAGATACAGCTGAAGGACAGCTCAAGAGACAGAGAGAGCAGCA 840
 QY 841 TTTCTAGCCAGTAAATGGCAATCATAGCAACATTCGATCATGCGCAGGTATTTAAT 900
 Db 841 TTTCTAGCCAGTAAATGGCAATCATAGCAACATTCGATCATGCGCAGGTATTTAAT 900
 QY 901 TTATGTAAACCTGGAAATCTGGGATCCAGTCTTAATAGGAGTACTTTGCAATACCAAT 960
 Db 901 TTATGTAAACCTGGAAATCTGGGATCCAGTCTTAATAGGAGTACTTTGCAATACCAAT 960
 QY 961 ATGGAATAAGGCGAGCTACTTGAAGTCTTTATGATATATTTGCTCTTCTCTACCT 1020
 Db 961 ATGGAATAAGGCGAGCTACTTGAAGTCTTTATGATATATTTGCTCTTCTCTACCT 1020
 QY 1021 GTTGTGACTGAGGAGTTTCATAGAGCACTACTCAGTGTAGATCCAGGAGGTTCCAAAGAC 1080
 Db 1021 GTTGTGACTGAGGAGTTTCATAGAGCACTACTCAGTGTAGATCCAGGAGGTTCCAAAGAC 1080
 QY 1081 AGTTGGAGGCTTTGAGATGCTTTGTCAGCTGAGGCAAAAGTATTTCTCTCATCGT 1140
 Db 1081 AGTTGGAGGCTTTGAGATGCTTTGTCAGCTGAGGCAAAAGTATTTCTCTCATCGT 1140
 QY 1141 GCCAGATCCAGGCGAGCTCATGATATTTATTTGCACTGATCTCTGCAATTTAT 1200
 Db 1141 GCCAGATCCAGGCGAGCTCATGATATTTATTTGCACTGATCTCTGCAATTTAT 1200
 QY 1201 CGTAATGGACTTTTATAGAGGCTAGTTGAAGTGATGATAAACAAGTATGATCATATCTCA 1260

Db 1201 CGTAATGGACTTTTATAGAGGCTAGTTGAAGTGATGATAAACAAGTATGATCATATCTCA 1260
 QY 1261 GTTAGAGCTACCATCCCTTTTAGAGAGCTTTTACATATGAGCAAAACAATTTCTTCTCAT 1320
 Db 1261 GTTAGAGCTACCATCCCTTTTAGAGAGCTTTTACATATGAGCAAAACAATTTCTTCTCAT 1320
 QY 1321 TCATATAGCCATCATTTACACTGCTTCCCAACCTTAATGAATATGCTGCTGATCCTTTGAT 1380
 Db 1321 TCATATAGCCATCATTTACACTGCTTCCCAACCTTAATGAATATGCTGCTGATCCTTTGAT 1380
 QY 1381 ATCCCCAAGCAAAAGAGACTGCGAGCCAGTGCAGCTTGAACCTTTTAAACCGCTTCCAT 1440
 Db 1381 ATCCCCAAGCAAAAGAGACTGCGAGCCAGTGCAGCTTGAACCTTTTAAACCGCTTCCAT 1440
 QY 1441 GAAATGAAGAAACGAGGAGCTTAAGCTTTATAGTCTTCAATTTAGACCAATATTTAGAAA 1500
 Db 1441 GAAATGAAGAAACGAGGAGCTTAAGCTTTATAGTCTTCAATTTAGACCAATATTTAGAAA 1500
 QY 1501 GCAATTCGAACACACAGGAGGATCAGTATCTCGAGTTCAGAAAGATATATTTATC 1560
 Db 1501 GCAATTCGAACACACAGGAGGATCAGTATCTCGAGTTCAGAAAGATATATTTATC 1560
 QY 1561 CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGTCTTCAACAT 1620
 Db 1561 CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAAAGTCTTCAACAT 1620
 QY 1621 AAAGAGAACTTTGAATGGAATCTTATAGGACCATTTCTTAAGTGGCCAAATGTA 1680
 Db 1621 AAAGAGAACTTTGAATGGAATCTTATAGGACCATTTCTTAAGTGGCCAAATGTA 1680
 QY 1681 AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGATCTATTTATTTT 1740
 Db 1681 AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGATCTATTTATTTT 1740
 QY 1741 TACAAGCCAGCAGTAAATATATATGCGCAACCTGGATCTGGATTTGCGCAAGGCCAAGAC 1800
 Db 1741 TACAAGCCAGCAGTAAATATATATGCGCAACCTGGATCTGGATTTGCGCAAGGCCAAGAC 1800
 QY 1801 CTCAGGTTGTAGTGTGCGAGTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGGCAA 1860
 Db 1801 CTCAGGTTGTAGTGTGCGAGTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGGCAA 1860
 QY 1861 GGCTACTTAGAAGATCTAGTAAAGGATATGTTCAAGTGGCTCAATGCTTCAATCTGGAATG 1920
 Db 1861 GGCTACTTAGAAGATCTAGTAAAGGATATGTTCAAGTGGCTCAATGCTTCAATCTGGAATG 1920
 QY 1921 AAACCCGAAGAGTCTTCAAAATTAATGTTTATGACCCCTTAGTCAACACTACTTTT 1980
 Db 1921 AAACCCGAAGAGTCTTCAAAATTAATGTTTATGACCCCTTAGTCAACACTACTTTT 1980
 QY 1981 TTATTTATTTGGAACACTTTTCTTCCACCTCATGAGGTTAAATGCTCGAAAAATGCAAGT 2040
 Db 1981 TTATTTATTTGGAACACTTTTCTTCCACCTCATGAGGTTAAATGCTCGAAAAATGCAAGT 2040
 QY 2041 GTATTTCAAGTGTCTTCAAAATTAATGTTTATGAGGATGATGATGATGATGATGATGATG 2100
 Db 2041 GTATTTCAAGTGTCTTCAAAATTAATGTTTATGAGGATGATGATGATGATGATGATGATG 2100
 QY 2101 ACTGTTCTTAGCTTGCAGCTATAGCAGAGATGATGATGATGATGATGATGATGATGATG 2160
 Db 2101 ACTGTTCTTAGCTTGCAGCTATAGCAGAGATGATGATGATGATGATGATGATGATGATG 2160
 QY 2161 TTAACCTGACGCTACTGATGCTTGCAGACTCTATGCAACAAACATTTAAGGGATATTTG 2220
 Db 2161 TTAACCTGACGCTACTGATGCTTGCAGACTCTATGCAACAAACATTTAAGGGATATTTG 2220
 QY 2221 AGAGCTAATGTTGAATTTCTTAAATTAATGTTGGGAAATGTTAGTGTAGTCCAGCTACAT 2280
 Db 2221 AGAGCTAATGTTGAATTTCTTAAATTAATGTTGGGAAATGTTAGTGTAGTCCAGCTACAT 2280
 QY 2281 GATAAAAAAACAACGATTTCTCTGAAAGCTCTTGTATATCTCTGATGAAGCATGTGAAGAC 2340

Db 2281 GATAAACAACAAACGATTTCCTCTGAGAGCTCTTGATATATCTCTGATGAAGCATGTGAAGAC 2340
Qy 2341 AAGGCAATCTTCATGCTCTCAATCAGATGAACACAGAGTTATCCCACTTTGGAGACAAG 2400
Db 2341 AAGGCAATCTTCATGCTCTCAATCAGATGAACACAGAGTTATCCCACTTTGGAGACAAG 2400
Qy 2401 GGTTCCTTCCTGCTGAGATTTCTCTCCATTCCAAAGGATTTTCTATCTGATGAA 2460
Db 2401 GGTTCCTTCCTGCTGAGATTTCTCTCCATTCCAAAGGATTTTCTATCTGATGAA 2460
Qy 2461 AGAGGTATGTAGCAAAAACATTTGAAAGTGGCACAGGGAATACAACTCCCAATATGTT 2520
Db 2461 AGAGGTATGTAGCAAAAACATTTGAAAGTGGCACAGGGAATACAACTCCCAATATGTT 2520
Qy 2521 GACTTGATTTGAGGAAACAATCAATGAAGCATTACTACTACCGGAAGCCTGTTGATGGT 2580
Db 2521 GACTTGATTTGAGGAAACAATCAATGAAGCATTACTACTACCGGAAGCCTGTTGATGGT 2580
Qy 2581 GATAACTATGTCGCGAGTAACCAAAAGATTAAGCGTCTCAGCTCTCAGCTCTACTGCTATA 2640
Db 2581 GATAACTATGTCGCGAGTAACCAAAAGATTAAGCGTCTCAGCTCTCAGCTCTACTGCTATA 2640
Qy 2641 CACCTTTATGACAACTAGTACACCATAAACAGGCTGCCATTTGTTGGAGTACAGAT 2700
Db 2641 CACCTTTATGACAACTAGTACACCATAAACAGGCTGCCATTTGTTGGAGTACAGAT 2700
Qy 2701 ATTATTAAGAACTCTGTCGTAATGTCGTACACAGATTTGGATAAGTGGGAAGAAATT 2760
Db 2701 ATTATTAAGAACTCTGTCGTAATGTCGTACACAGATTTGGATAAGTGGGAAGAAATT 2760
Qy 2761 AAAAATCTGAAGCATCTCTTTGGGCCCTTTGGGAAATATCGGCTCATCAAAATGGGGTCTC 2820
Db 2761 AAAAATCTGAAGCATCTCTTTGGGCCCTTTGGGAAATATCGGCTCATCAAAATGGGGTCTC 2820
Qy 2821 AATTGCTACAGGAGAAACGTAATCCAGATATCTAAACTTGAACACTTGAACACTGTTGAA 2880
Db 2821 AATTGCTACAGGAGAAACGTAATCCAGATATCTAAACTTGAACACTTGAACACTGTTGAA 2880
Qy 2881 GTTCTTTCCATCAGAGGACCTGTGTATATGTATGTACTTTGGGCTCATAGCTTAAACCAAACAA 2940
Db 2881 GTTCTTTCCATCAGAGGACCTGTGTATATGTATGTACTTTGGGCTCATAGCTTAAACCAAACAA 2940
Qy 2941 GGCTGTGATATCTAAATGTCAAACTGGGATGCTGTAGGCAATGTCGCAACATCTG 3000
Db 2941 GGCTGTGATATCTAAATGTCAAACTGGGATGCTGTAGGCAATGTCGCAACATCTG 3000
Qy 3001 TGGCCAGTGGTTCAGATGATGTGGAACAACCTCTGTATGAACCTTTCATCTATCCCAAGC 3060
Db 3001 TGGCCAGTGGTTCAGATGATGTGGAACAACCTCTGTATGAACCTTTCATCTATCCCAAGC 3060
Qy 3061 ACTCTAAGTTTGAACCTCGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTGTG 3120
Db 3061 ACTCTAAGTTTGAACCTCGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTGTG 3120
Qy 3121 CCATCGAGTATGTTCAATTCGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
Db 3121 CCATCGAGTATGTTCAATTCGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
Qy 3181 TTCCTTGATATCAATGAAGATACAGAGCAACATTTATGACCGATCTGGACCCATAAG 3240
Db 3181 TTCCTTGATATCAATGAAGATACAGAGCAACATTTATGACCGATCTGGACCCATAAG 3240
Qy 3241 GATAAATTCATTTCCCTCTCTTCTCTAGTAACTTTGGAAGATCGTATCTTAAAT 3300
Db 3241 GATAAATTCATTTCCCTCTCTTCTCTAGTAACTTTGGAAGATCGTATCTTAAAT 3300
Qy 3301 TCGCTTACTTTGCCCTAAACAAAACATCGTAGTACAGTATCCAAAGGAGGGAATTA 3360
Db 3301 TCGCTTACTTTGCCCTAAACAAAACATCGTAGTACAGTATCCAAAGGAGGGAATTA 3360
Qy 3361 TCATCTGAAGTAAAGCAAGCAACAGGCGAATACAGACATTTACGAGCCCAAGTGTGAT 3420
Db 3361 TCATCTGAAGTAAAGCAAGCAACAGGCGAATACAGACATTTACGAGCCCAAGTGTGAT 3420

Qy 3421 TTTAATCATAGTATGATTTTACCCATATCCACTGTACAGAAAACATTACAAATTAGAG 3480
Db 3421 TTTAATCATAGTATGATTTTACCCATATCCACTGTACAGAAAACATTACAAATTAGAG 3480
Qy 3481 ACTTCATTTATGGGGAATAAGCAATTTGAAGACATCTGGTAGTACACCAAGCATTTGAGAA 3540
Db 3481 ACTTCATTTATGGGGAATAAGCAATTTGAAGACATCTGGTAGTACACCAAGCATTTGAGAA 3540
Qy 3541 AATGACTTTAAATTTCAACCAAGAAATTTTGGTACAGAGAAATACAGAGAAATACAGCCGA 3600
Db 3541 AATGACTTTAAATTTCAACCAAGAAATTTTGGTACAGAGAAATACAGAGAAATACAGCCGA 3600
Qy 3601 GAGAGTTAGTAGTAAAGTTTCAACGAGCTCACATATGAAGATACGTAGTCCAAAGTTTC 3660
Db 3601 GAGAGTTAGTAGTAAAGTTTCAACGAGCTCACATATGAAGATACGTAGTCCAAAGTTTC 3660
Qy 3661 AATACAGACACTACAAAGTGGCATTAAGTTTCAATGAGCTCAAGTCTTTCACGAGAGACA 3720
Db 3661 AATACAGACACTACAAAGTGGCATTAAGTTTCAATGAGCTCAAGTCTTTCACGAGAGACA 3720
Qy 3721 GTAGGTGTAGATGCTACAACTATGACACAGACTGTGGAGCATGTGTGTAAGT 3780
Db 3721 GTAGGTGTAGATGCTACAACTATGACACAGACTGTGGAGCATGTGTGTAAGT 3780
Qy 3781 ACTAAAACTATTAAGACAAAGCCACTATTTGACGCCACAGTCTTAACCATCTCTCTCTCC 3840
Db 3781 ACTAAAACTATTAAGACAAAGCCACTATTTGACGCCACAGTCTTAACCATCTCTCTCTCC 3840
Qy 3841 AAATCAAAATTCGGTGTCCCTGGTCCCTCAAGTTCTTCTCATAGCTTCTCTAGAGAGCA 3900
Db 3841 AAATCAAAATTCGGTGTCCCTGGTCCCTCAAGTTCTTCTCATAGCTTCTCTAGAGAGCA 3900
Qy 3901 CAGTCCCTTAAAGCACCCCTCTATTTGCTACAAATTTAAAGTCTAGCAGATTTGTAATTTAGT 3960
Db 3901 CAGTCCCTTAAAGCACCCCTCTATTTGCTACAAATTTAAAGTCTAGCAGATTTGTAATTTAGT 3960
Qy 3961 TACACAAAGTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAAGA 4020
Db 3961 TACACAAAGTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAAGA 4020
Qy 4021 ATGATCCATCTCTATCTCACTCTGAAGCTTGGCATCTCCAGCAAAAGATGTGCTATTT 4080
Db 4021 ATGATCCATCTCTATCTCACTCTGAAGCTTGGCATCTCCAGCAAAAGATGTGCTATTT 4080
Qy 4081 ACTGATACCATCAACATGAAGGCCAACAGTTTGGAGTCCAGATTTAAACCAAGCAGGTTTC 4140
Db 4081 ACTGATACCATCAACATGAAGGCCAACAGTTTGGAGTCCAGATTTAAACCAAGCAGGTTTC 4140
Qy 4141 ATGAAGCCTTAAAGTTATGATCATTAATAGATAAAGAAATTTATTGAGTCTCTATTAAATCAA 4200
Db 4141 ATGAAGCCTTAAAGTTATGATCATTAATAGATAAAGAAATTTATTGAGTCTCTATTAAATCAA 4200
Qy 4201 AATAACCTGCAACGATCTTCTCAGTCCGCTCCATGCTGCTCAGTCCACATATGGGGGT 4260
Db 4201 AATAACCTGCAACGATCTTCTCAGTCCGCTCCATGCTGCTCAGTCCACATATGGGGGT 4260
Qy 4261 TCAGATCATTTACATTTGCTCTCCCGTGGATATAAATGATATATTCCAGGTAAG 4320
Db 4261 TCAGATCATTTACATTTGCTCTCCCGTGGATATAAATGATATATTCCAGGTAAG 4320
Qy 4321 GATATTCCTTATTTTCAGACAAAACATACACCATGATGATCGAGGTGCAAGAGCA 4380
Db 4321 GATATTCCTTATTTTCAGACAAAACATACACCATGATGATCGAGGTGCAAGAGCA 4380
Qy 4381 TTTGCCCATGATGCGAGGCTCTTCCATCTGGAACTGGAGTCTTTGTAATAAAATTTCTTTT 4440
Db 4381 TTTGCCCATGATGCGAGGCTCTTCCATCTGGAACTGGAGTCTTTGTAATAAAATTTCTTTT 4440
Qy 4441 CACTTGTCTAGCAGCAGATGAGTCTTACGGAATAATGAAATTCATTCATTCAGATGCC 4500
Db 4441 CACTTGTCTAGCAGCAGATGAGTCTTACGGAATAATGAAATTCATTCATTCAGATGCC 4500

QY	4501	TCCTCTGTTTTAGAAAGTACAGAACACCTGACCTACAGGAAACATACAGATGATAACTGC	4560	Db	28	ATGGCGCGATCGCGCGCGCGCTCTCTGAAAGAACCTCCGAGTACAGGCGGGAATGAC	87
Db	4501	TCCTCTGTTTTAGAAAGTACAGAACACCTGACCTACAGGAAACATACAGATGATAACTGC	4560	QY	61	AGCGGCGAGGAGAAACGTCCCGCTGGATCTGACCCGAGAACCTCTGATACCTTAAGAGAG	120
QY	4561	CTTTATTTGTCTGTATTGAATTCCTGAGTTTCCAGCCAGCAACCAACTGAGTGAATA	4620	Db	88	AGCGGCGAGGAGAAACGTCCCGCTGGATCTGACCCGAGAACCTCTGATACCTTAAGAGAG	147
Db	4561	CTTTATTTGTCTGTATTGAATTCCTGAGTTTCCAGCCAGCAACCAACTGAGTGAATA	4620	QY	121	ATTCTCCAAATCTGGCCAGATTGCGAGGAGTATCAATATATGAGAAAGCTAGGCCATCTG	180
QY	4621	TGTAGTCATTACAGACTTTCCAGATATTCCTGATTTCTGATTTGTTGAGCAGACTATCCAT	4680	Db	148	ATTCTCCAAATCTGGCCAGATTGCGAGGAGTATCAATATATGAGAAAGCTAGGCCATCTG	207
Db	4621	TGTAGTCATTACAGACTTTCCAGATATTCCTGATTTCTGATTTGTTGAGCAGACTATCCAT	4680	QY	181	AATAACTTTTACTAAGCTTTCTTTGATATATGCGCCACAGTGAAGAAAACCTGGGCTTTCAC	240
QY	4681	AATCCCTTTAGAAAGTGTCCCTTAAAGTTTTCGGGATTTTCGGATGCGATGCGGGTG	4740	Db	208	AATAACTTTTACTAAGCTTTCTTTGATATATGCGCCACAGTGAAGAAAACCTGGGCTTTCAC	267
Db	4681	AATCCCTTTAGAAAGTGTCCCTTAAAGTTTTCGGGATTTTCGGATGCGATGCGGGTG	4740	QY	241	TATGAGGATATCATAAATTTCTTTGCGGTAGCTTTTATTAATGAAGCAAAAGAAAGTGCGA	300
QY	4741	TCTCAAGAGGCTCAGCTAGCAGCAACCAAGACACAGAAATTTACTAGGTGTAAACA	4800	Db	268	TATGAGGATATCATAAATTTCTTTGCGGTAGCTTTTATTAATGAAGCAAAAGAAAGTGCGA	327
Db	4741	TCTCAAGAGGCTCAGCTAGCAGCAACCAAGACACAGAAATTTACTAGGTGTAAACA	4800	QY	301	GCAGCAGGCTTACGAGCGCTTCGATATCTCATCCAGACTCCAGATATTTCCAGAGGTTG	360
QY	4801	ATTCCAGATACACCAATGTCCTGATATCTCTTCGCAAGAAATTTCTAAGATTAGTC	4860	Db	328	GCAGCAGGCTTACGAGCGCTTCGATATCTCATCCAGACTCCAGATATTTCCAGAGGTTG	387
Db	4801	ATTCCAGATACACCAATGTCCTGATATCTCTTCGCAAGAAATTTCTAAGATTAGTC	4860	QY	361	CTAAAAATTGAAGTGGACTATTTTAATAGCTAGGTGCATTCGACATACAAAGCAAGCGAG	420
QY	4861	ATTAAATTGAGTGTGAGTTTCAATTAATGTCAGACTGGGCTTTTAAACAATTAG	4920	Db	388	CTAAAAATTGAAGTGGACTATTTTAATAGCTAGGTGCATTCGACATACAAAGCAAGCGAG	447
Db	4861	ATTAAATTGAGTGTGAGTTTCAATTAATGTCAGACTGGGCTTTTAAACAATTAG	4920	QY	421	GTAGAGGACACAAAGCACTTCGATTTAGTCAAGAAAGATGATTAATCTGTAATGCTCTTCTG	480
QY	4921	GAGNAGTATCTCAACATTTGATGACATATGCTTTACTCTGAGGTTCCTCAATTTGCTG	4980	Db	448	GTAGAGGACACAAAGCACTTCGATTTAGTCAAGAAAGATGATTAATCTGTAATGCTCTTCTG	507
Db	4921	GAGNAGTATCTCAACATTTGATGACATATGCTTTACTCTGAGGTTCCTCAATTTGCTG	4980	QY	481	TTTCTCTAGTCTGTGACCAACTCATTAATTTGAGTTGGAATGATGACTTCAGAAAGA	540
QY	4981	TCACACTGCACATTCAGACTTCGCTGTCGAGGTTTCATCAAGAAATTTTCAAGATGA	5040	Db	508	TTTCTCTAGTCTGTGACCAACTCATTAATTTGAGTTGGAATGATGACTTCAGAAAGA	567
Db	4981	TCACACTGCACATTCAGACTTCGCTGTCGAGGTTTCATCAAGAAATTTTCAAGATGA	5040	QY	541	GACAGAAATGCTCCGAGCATGCTTGCCTATTTCTGTAATCTGACATTCAGAAATCCAGAG	600
QY	5041	CAGTTTCTCAAAATGCAATGCAAGGCTGTTGTCGAGGTTTCATCAAGAAATTTTCAAGATGA	5100	Db	568	GACAGAAATGCTCCGAGCATGCTTGCCTATTTCTGTAATCTGACATTCAGAAATCCAGAG	627
Db	5041	CAGTTTCTCAAAATGCAATGCAAGGCTGTTGTCGAGGTTTCATCAAGAAATTTTCAAGATGA	5100	QY	601	GTGTTGGGCTTCGAGGAGGACTAAACCAATATTTGAAAAATGATGATGATGCAATTA	660
QY	5101	ATAGTTGATACATCTGCTGAATCTGA	5127	Db	628	GTGTTGGGCTTCGAGGAGGACTAAACCAATATTTGAAAAATGATGATGATGCAATTA	687
Db	5101	ATAGTTGATACATCTGCTGAATCTGA	5127	QY	661	AGTCGAATAATGAGGCGCTTAATTTACTACAAATTTTGGACCTTTCTTAATCATCCAAAGACT	720
RESULT 2							
US-10-782-244-1							
; Sequence 1, Application US/10782244							
; Publication No. US20040253677A1							
; GENERAL INFORMATION:							
; APPLICANT: Sabatini, David M. ✓							
; APPLICANT: Kim, Do-Hyung							
; APPLICANT: Saitobay, Dos D.							
; TITLE OF INVENTION: MTOR KINASE-ASSOCIATED PROTEINS							
; FILE REFERENCE: WIPO-01-010							
; CURRENT APPLICATION NUMBER: US/10/782,244							
; PRIOR FILING DATE: 2004-02-18							
; PRIOR APPLICATION NUMBER: US 60/448,035							
; NUMBER OF SEQ ID NOS: 51							
; SOFTWARE: FastSeq for Windows Version 4.0							
; SEQ ID NO 1							
; LENGTH: 5662							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
US-10-782-244-1							
Query Match 100.0%; Score 5127; DB 18; Length 5662;							
Best Local Similarity 100.0%; Pred. No. 0;							
Matches 5127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY	1	ATGGCGCGATCGCGCGCGCTCTCTGAAAGAACCTCCGAGTACAGGCGGGAATGAC	60				

1141 GCCAGATCCAGGCCAGACCTCATGGATAATTAATTTGGCACTGATACTCTCTGCAATTAAT 1200
1168 GCCAGATCCAGGCCAGACCTCATGGATAATTAATTTGGCACTGATACTCTCTGCAATTAAT 1227
1201 CGTAATGACATTTAGAGGGCTAGTTCGAAGTGAATAACAAACAGTGATGATCATCTCA 1260
1228 CGTAATGACATTTAGAGGGCTAGTTCGAAGTGAATAACAAACAGTGATGATCATCTCA 1287
1261 GTTAGAGCTACCATCTTTTATAGGAGAGCTTTTACATATGGCAACACAAATTTCTTCCAT 1320
1288 GTTAGAGCTACCATCTTTTATAGGAGAGCTTTTACATATGGCAACACAAATTTCTTCCAT 1347
1321 TCACATAGCCATCAATTTACATGCTTGCCAAACCCCTAATGAATATAGGCTGCACTCTTTGAT 1380
1348 TCACATAGCCATCAATTTACATGCTTGCCAAACCCCTAATGAATATAGGCTGCACTCTTTGAT 1407
1381 ATCCCCAAGGAAGAGAGCTGGAGCCAGTCGAGCCTTGAACCTTTTAAACGCTTCCAT 1440
1408 ATCCCCAAGGAAGAGAGCTGGAGCCAGTCGAGCCTTGAACCTTTTAAACGCTTCCAT 1467
1441 GAAATGAAGAAACGAGGACCTAAGCCTTATAGTCTTCAATTTAGACCACATTAATTCAGAAA 1500
1468 GAATGAAGAAACGAGGACCTAAGCCTTATAGTCTTCAATTTAGACCACATTAATTCAGAAA 1527
1501 GCAATTCGAACACACACGAGAAACGGGATCAGTATCTCCGAGTTCAGAAAGATATATTTATC 1560
1528 GCAATTCGAACACACGAGAAACGGGATCAGTATCTCCGAGTTCAGAAAGATATATTTATC 1587
1561 CTTAAGGATCAGAGGAAGCTCTTTAATTAACCTTAGAGATAGCAAGTCCCTTCAACAT 1620
1588 CTTAAGGATCAGAGGAAGCTCTTTAATTAACCTTAGAGATAGCAAGTCCCTTCAACAT 1647
1621 AAAGAGAACTTTGAATGGAACTTATATAGGACCATTTTAAGTGGCCAAATGTA 1680
1648 AAAGAGAACTTTGAATGGAACTTATATAGGACCATTTTAAGTGGCCAAATGTA 1707
1681 AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT 1740
1708 AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT 1767
1741 TACAAGCCAGCAGTAATTAATATGCCAACCTGATCTGGATTTTGGCAAGGCCAAACAG 1800
1768 TACAAGCCAGCAGTAATTAATATGCCAACCTGATCTGGATTTTGGCAAGGCCAAACAG 1827
1801 CTACAGGTTGTAGGTTGCGAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGCAA 1860
1828 CTACAGGTTGTAGGTTGCGAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGCAA 1887
1861 GGCTACTTAGAAGATCTAGTAAAGGATATTTGTTCAAGTGGCTCAATGCTTCACTGGAAATG 1920
1888 GGCTACTTAGAAGATCTAGTAAAGGATATTTGTTCAAGTGGCTCAATGCTTCACTGGAAATG 1947
1921 AAACCCGAAGAAGCTCTTCAAAATAATGGTTTATATGACCAACCTTAGTCAACACTATTT 1980
1948 AAACCCGAAGAAGCTCTTCAAAATAATGGTTTATATGACCAACCTTAGTCAACACTATTT 2007
1981 TTATTTATGGAACACTTTCTTGGCACCCCTCATCGAGTTTAAATGCTGGAAAATGCGAGT 2040
2008 TTATTTATGGAACACTTTCTTGGCACCCCTCATCGAGTTTAAATGCTGGAAAATGCGAGT 2067
2041 GTATTTCAAGTGTCTCTTAATCTTTGCTCTTGGAAAACCAAGATCACTTGTCTAAAACCTT 2100
2068 GTATTTCAAGTGTCTCTTAATCTTTGCTCTTGGAAAACCAAGATCACTTGTCTAAAACCTT 2127
2101 ACTGTTCTAGCTGGACTATAGCAGAGATGGATGGGTGAGTGCATCTTCCAAATTT 2160
2128 ACTGTTCTAGCTGGACTATAGCAGAGATGGATGGGTGAGTGCATCTTCCAAATTT 2187
2161 TTAATCGAGCTAGTGCCTGAGACTCTATGCAACAAACATTTTAAGGGTATTTATG 2220
2188 TTAATCGAGCTAGTGCCTGAGACTCTATGCAACAAACATTTTAAGGGTATTTATG 2247

2221 AGAGCTAATGTTGAATCTTTTAATAATTTGGGAAATGAGTTGTTAGTACCCAGCTACAT 2280
2248 AGAGCTAATGTTGAATCTTTTAATAATTTGGGAAATGAGTTGTTAGTACCCAGCTACAT 2307
2281 GATAAAAACAAAACGATTTCTCTGAAGCTCTTGATATCTCGATGAAGCATGTGAAGAC 2340
2308 GATAAAAACAAAACGATTTCTCTGAAGCTCTTGATATCTCGATGAAGCATGTGAAGAC 2367
2341 AAGGCCAATCTTCAATGCTCTCATTCAGATGAACACAGGTTATCCCACTTTGGAGACAAG 2400
2368 AAGGCCAATCTTCAATGCTCTCATTCAGATGAACACAGGTTATCCCACTTTGGAGACAAG 2427
2401 GGTTTGTCTTCTCTGCTGAGATTTCTCTCCATTTCCAAAAGGATTTCTCTATCTGAATGAA 2460
2428 GGTTTGTCTTCTCTGCTGAGATTTCTCTCCATTTCCAAAAGGATTTCTCTATCTGAATGAA 2487
2461 AGAGGTTATGTAGCAAAAACAAATTTGGAAAAGTGGCACAGGGAATACAACCTCCAAATATGTT 2520
2488 AGAGGTTATGTAGCAAAAACAAATTTGGAAAAGTGGCACAGGGAATACAACCTCCAAATATGTT 2547
2521 GACTTGATTCGAGGAACAACTCAATGAAGCACTTACTACTTACCGGAAGCCCTGTTGATGGT 2580
2548 GACTTGATTCGAGGAACAACTCAATGAAGCACTTACTACTTACCGGAAGCCCTGTTGATGGT 2607
2581 GATAAATCTATGTTGTCGAGTAACCAAAGATTAACAGCTCTCTCACTGCTACCTGCTCTATA 2640
2608 GATAAATCTATGTTGTCGAGTAACCAAAGATTAACAGCTCTCTCACTGCTACCTGCTCTATA 2667
2641 CACCTTTATGGAACAACTAGTACACCATAAAACAGGCTGCCATTTGTTGGAAGTACAGAAAT 2700
2668 CACCTTTATGGAACAACTAGTACACCATAAAACAGGCTGCCATTTGTTGGAAGTACAGAAAT 2727
2701 ATTATTACAGAACTCTCTGCTGAATGTTGCTACACAGATTTGGAATAGTGGGAAGAAAT 2760
2728 ATTATTACAGAACTCTCTGCTGAATGTTGCTACACAGATTTGGAATAGTGGGAAGAAAT 2787
2761 AAAAACTGAAAGCATCTCTTTGGGCTTTGGGAAATATCGGCTCATCAAAATTTGGGGTCTC 2820
2788 AAAAACTGAAAGCATCTCTTTGGGCTTTGGGAAATATCGGCTCATCAAAATTTGGGGTCTC 2847
2821 AATTTGCTACAGGAAGAAACGTTGTCAGATATATCTAAAACTTGGCAAAACAGTGTGAA 2880
2848 AATTTGCTACAGGAAGAAACGTTGTCAGATATATCTAAAACTTGGCAAAACAGTGTGAA 2907
2881 GTTCTTTCCATCAGAGGACCTGTGTATATGTAATCTGGGCTCATAGCTTAAAAACCAACAA 2940
2908 GTTCTTTCCATCAGAGGACCTGTGTATATGTAATCTGGGCTCATAGCTTAAAAACCAACAA 2967
2941 GGCTGTGATATTTCTAAAACTCAACCTGGGATGCTGTGAGGCATAGTCGCAAAACATCTG 3000
2968 GGCTGTGATATTTCTAAAACTCAACCTGGGATGCTGTGAGGCATAGTCGCAAAACATCTG 3027
3001 TGGCCAGTGTTCAGATGATGTGGAAACAACTCTGTGAATGAACTTTTCACTATCTCCCAAGC 3060
3028 TGGCCAGTGTTCAGATGATGTGGAAACAACTCTGTGAATGAACTTTTCACTATCTCCCAAGC 3087
3061 ACTCTAAGTTTGAACCTCGAGTCAACCCAGCTCTAGACATATAGTGAAGATGAAATCTGTG 3120
3088 ACTCTAAGTTTGAACCTCGAGTCAACCCAGCTCTAGACATATAGTGAAGATGAAATCTGTG 3147
3121 CCATCGAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
3148 CCATCGAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3207
3181 TTCTTTGATATCAATGAAGATACAGAGCAACATTTTATGACCGATCTGGACCCATATAAG 3240
3208 TTCTTTGATATCAATGAAGATACAGAGCAACATTTTATGACCGATCTGGACCCATATAAG 3267
3241 GATAAAAAATTCATTTCCCTTTCTTTGCTTTAGTAAACTTGTGAAGATCGTATCTTAAAT 3300
3268 GATAAAAAATTCATTTCCCTTTCTTTGCTTTAGTAAACTTGTGAAGATCGTATCTTAAAT 3327
3301 TCGCTTACTTTGCTTAACAAAAAACATCTGTAAGTGTGATCCAAAGGGGAAATTA 3360

Db 3328 TCGCTTACTTTGCCCTAAACAAAACATCGTAGTAGCAGTGATCCAAAAGAGGGGAATTA 3387
QY 3361 TCATCTGAAAGTAAGAACAAGCAACAGCGGAATCAGAACATTTACGAGCCCAAGTGTGAT 3420
Db 3388 TCATCTGAAAGTAAGAACAAGCAACAGCGGAATCAGAACATTTACGAGCCCAAGTGTGAT 3447
QY 3421 TTTAATCATAGTAGATGATTTTACACCCATATCCACTGTACAGAAAACATTAACATTAAG 3480
Db 3448 TTTAATCATAGTAGATGATTTTACACCCATATCCACTGTACAGAAAACATTAACATTAAG 3507
QY 3481 ACTTCATTTATCGGGAAATAGACACATTTGAAGACACTGGTAGTACACCAAGCAATTCGAGAA 3540
Db 3508 ACTTCATTTATCGGGAAATAGACACATTTGAAGACACTGGTAGTACACCAAGCAATTCGAGAA 3567
QY 3541 AATGACTTAAATTTCAACCAAGAAATTTGGTAAGAGAAATCAAGAGCGGA 3600
Db 3568 AATGACTTAAATTTCAACCAAGAAATTTGGTAAGAGAAATCAAGAGCGGA 3627
QY 3601 GAGAGGTAGTAGTAAAGTTCAACGAGCTCACATATGAAGATAGGTAGGCCAAAGTTTC 3660
Db 3628 GAGAGGTAGTAGTAAAGTTCAACGAGCTCACATATGAAGATAGGTAGGCCAAAGTTTC 3687
QY 3661 AATACAGACACTCAACCAAGTGCATAAGTTCAATGAGCTCAAGTCTTTCACGAGAGACA 3720
Db 3688 AATACAGACACTCAACCAAGTGCATAAGTTCAATGAGCTCAAGTCTTTCACGAGAGACA 3747
QY 3721 GTAGGTGTAGATGCTCAACTATPGBACACAGACTGTGGAAGCAGTAGTACTGTGTAAGT 3780
Db 3748 GTAGGTGTAGATGCTCAACTATPGBACACAGACTGTGGAAGCAGTAGTACTGTGTAAGT 3807
QY 3781 ACTAATACTATTAAAGCAAGCCACTATTGACGCCACAGCTAACCATCTGCTCTCTCC 3840
Db 3808 ACTAATACTATTAAAGCAAGCCACTATTGACGCCACAGCTAACCATCTGCTCTCTCC 3867
QY 3841 AAATCAAAATTCGGTGTCCCTGTCCTCCAGGTTCTTCTCATACGCTTCTTCTAGAGAGCA 3900
Db 3868 AAATCAAAATTCGGTGTCCCTGTCCTCCAGGTTCTTCTCATACGCTTCTTCTAGAGAGCA 3927
QY 3901 CAGTCCCTTAAAGACCTCTATGTCTAATTAAGTCTAGCAGATGTTAATCTTACT 3960
Db 3928 CAGTCCCTTAAAGACCTCTATGTCTAATTAAGTCTAGCAGATGTTAATCTTACT 3987
QY 3961 TACACAGTGTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAAGA 4020
Db 3988 TACACAGTGTCTAGAGATGCTTTTGGCTATGCTACACTGAAAAGACTACAGCAACAAAGA 4047
QY 4021 ATGCATCCATCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATT 4080
Db 4048 ATGCATCCATCCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGTCTATT 4107
QY 4081 ACTGATACCATCACCATGAAGGCCAACAGTTTGGTCCAGATTAACCAACGAGGTT 4140
Db 4108 ACTGATACCATCACCATGAAGGCCAACAGTTTGGTCCAGATTAACCAACGAGGTT 4167
QY 4141 ATGAAGCCTTAAGTATGATCATATTAGATAAGAGATTTATGAGTCTCTATTATCAA 4200
Db 4168 ATGAAGCCTTAAGTATGATCATATTAGATAAGAGATTTATGAGTCTCTATTATCAA 4227
QY 4201 AATACCTGCAACGATCTTCTCAGTGGGTCCATGTGTCAGTGCACATATGGGGT 4260
Db 4228 AATACCTGCAACGATCTTCTCAGTGGGTCCATGTGTCAGTGCACATATGGGGT 4287
QY 4261 TCAGATGATTAATGCTTGTCTCCCGTGTGATATAATGATATATTCCAGGTTAAAG 4320
Db 4288 TCAGATGATTAATGCTTGTCTCCCGTGTGATATAATGATATATTCCAGGTTAAAG 4347
QY 4321 GATATTCCTTATTTTCAGACAAAAACATACCAACATGATGATCGAGGTGCAAGACA 4380
Db 4348 GATATTCCTTATTTTCAGACAAAAACATACCAACATGATGATCGAGGTGCAAGACA 4407
QY 4381 TTTGCCCATGATCGAGGAGTCTTCCATCTGGAAGTCTGTGTAATAAAATTCCTTT 4440

Db 4408 TTTGCCCATGATCGAGGAGTCTTCCATCTGGAACCTGGAGGTCTTGTAAAAAATTCCTTT 4467
QY 4441 CACTTGTCTACGACAGATGAGTCTTACGGAATAATGAATTAATCAATCCATTCAGATGCC 4500
Db 4468 CACTTGTCTACGACAGATGAGTCTTACGGAATAATGAATTAATCAATTCAGATGCC 4527
QY 4501 TCTCTGTTTTTAGAAAGTACAGAAAGACACTGGACTACAGGAACATACAGATGATACTGC 4560
Db 4528 TCTCTGTTTTTAGAAAGTACAGAAAGACACTGGACTACAGGAACATACAGATGATACTGC 4587
QY 4561 CTTTATTTGTCTGTATTTGAAATTTCTGGTTTCCAGCCAGCAACCACTGAGTGCAATA 4620
Db 4588 CTTTATTTGTCTGTATTTGAAATTTCTGGTTTCCAGCCAGCAACCACTGAGTGCAATA 4647
QY 4621 TGTAGTCAATTCAGACTTTTCAAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4680
Db 4648 TGTAGTCAATTCAGACTTTTCAAGATATTTCCATATTTCTGATTTGGTGTGAGCAGACTATCCAT 4707
QY 4681 AATCCTTTAGAAAGTGGTCCCTCTAAGTTTTCGGGGATTTTCTGATGAGTGGGGTG 4740
Db 4708 AATCCTTTAGAAAGTGGTCCCTCTAAGTTTTCGGGGATTTTCTGATGAGTGGGGTG 4767
QY 4741 TCTCAAGAGGCTCAGCTAGCAGCACCAAGACACAGAAATTTGTTACTAGGTGTTAAAAACA 4800
Db 4768 TCTCAAGAGGCTCAGCTAGCAGCACCAAGACACAGAAATTTGTTACTAGGTGTTAAAAACA 4827
QY 4801 ATTCAGATGATACACCAATGTGCGGTATCTCTTCGAAAGAAAGTTCTAAGATTAGTC 4860
Db 4828 ATTCAGATGATACACCAATGTGCGGTATCTCTTCGAAAGAAAGTTCTAAGATTAGTC 4887
QY 4861 ATTAATTTGAGTAGTTCAAGTTTCACTAATGTATGAGACTGGCTTTTAAACAATTAAG 4920
Db 4888 ATTAATTTGAGTAGTTCAAGTTTCACTAATGTATGAGACTGGCTTTTAAACAATTAAG 4947
QY 4921 GAGAGTATCTCAAAACATTTGATGACATATGCTTACTCTGAGGTTTCCCATTTGCTG 4980
Db 4948 GAGAGTATCTCAAAACATTTGATGACATATGCTTACTCTGAGGTTTCCCATTTGCTG 5007
QY 4981 TCACACTGCACATTCAGACTTCCGTTGTCGAGGTTTCAACAAGAAATTAATTTCAAGATGTA 5040
Db 5008 TCACACTGCACATTCAGACTTCCGTTGTCGAGGTTTCAACAAGAAATTAATTTCAAGATGTA 5067
QY 5041 CAGTTTCTACAATGCATGAGAGCAGAGGCTGTTGGCAACACCAACCAAGCAACT 5100
Db 5068 CAGTTTCTACAATGCATGAGAGCAGAGGCTGTTGGCAACACCAACCAAGCAACT 5127
QY 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127
Db 5128 ATAGTTGATACATCTGCTGAATCCTGA 5154

RESULT 3
US-09-814-353-19068/c
; Sequence 19068, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: PabstSeq for Windows Version 4.0
; SEQ ID NO 19068
; LENGTH: 4657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 105, 4656, 4657
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19068

Query Match		34.6%;	Score 1772.8;	DB 10;	Length 4657;
Best Local Similarity		95.0%;	Pred. No. 0;		
Matches 1881;		Conservative	0;	Mismatches	2;
				Indels	97;
				Gaps	1;
Qy	3148	GACCGGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTGATATCAATGAAGATACAGAG	3207		
Db	4634	GACCGGTTTGGCAGCAGCTCTACTAGTACATTTTTCCTTGATATCAATGAAGATACAGAG	4575		
Qy	3208	CCAAACATTTTATGACCGATCTGGACCCCATAAAGATATAAAATTCATTCCTTTCTTTGCT	3267		
Db	4574	CCAAACATTTTATGACCGATCTGGACCCCATAAAGATATAAAATTCATTCCTTTCTTTGCT	4515		
Qy	3268	TCTAGTAACTTGTGAGAACTCGTATCTTAAATTCGCTTACTTTGCTTACCAAAAACAT	3327		
Db	4514	TCTAGTAACTTGTGAGAACTCGTATCTTAAATTCGCTTACTTTGCTTACCAAAAACAT	4455		
Qy	3328	CGTAGTAGCAGTGTATCCAAAGAGGAGGAAATTAATCATCTGAAAGTAAGACAAAGCAAGG	3387		
Db	4454	CGTAGTAGCAGTGTATCCAAAGAGGAGGAAATTAATCATCTGAAAGTAAGACAAAGCAAGG	4395		
Qy	3388	CGAATCAAGACACTTACCGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACCC	3447		
Db	4394	CGAATCAAGACACTTACCGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACACCC	4335		
Qy	3448	ATATCCACTGTACAGAAACATTTACAATTAGAGACTTTCATTTATGGGGAATAGCAATTT	3507		
Db	4334	ATATCCACTGTACAGAAACATTTACAATTAGAGACTTTCATTTATGGGGAATAGCAATTT	4275		
Qy	3508	GAAGACACTGGTAGTACACCAAGCATTTGGAGAAATGACTTTAAATTCACCAAGAAATTTT	3567		
Db	4274	GAAGACACTGGTAGTACACCAAGCATTTGGAGAAATGACTTTAAATTCACCAAGAAATTTT	4215		
Qy	3568	GGTACAGAGAAATCACAGAGAAATACAAAGCCGAGAGAGGTTAGTAGTAAAGTTCAACG	3627		
Db	4214	GGTACAGAGAAATCACAGAGAAATACAAAGCCGAGAGAGGTTAGTAGTAAAGTTCAACG	4155		
Qy	3628	AGCTCATATGAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATTA	3687		
Db	4154	AGCTCATATGAGATACGTAGCCAAAGTTTCAATACAGACACTACAAAGTGGCATTA	4095		
Qy	3688	AGTTCAATGAGCTCAAGTCTCTTACAGAGACAGTGTAGTGTGTACAACTATATGGAC	3747		
Db	4094	AGTTCAATGAGCTCAAGTCTCTTACAGAGACAGTGTAGTGTGTACAACTATATGGAC	4035		
Qy	3748	ACAGACTGTGGAAGCATGAGTACTGTGTAAAGTAACTATAAAGCAAGCCACTAT	3807		
Db	4034	ACAGACTGTGGAAGCATGAGTACTGTGTAAAGTAACTATAAAGCAAGCCACTAT	4009		
Qy	3808	TTGAGGCCACAGCTTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCT	3867		
Db	4008	TTGAGGCCACAGCTTAACCATCTGTCTCTCCAAATCAAAATTCGGTGTCCCTGGTGCCT	4009		
Qy	3868	CCAGGTTTCTTCTCATACGCTTCTTGAAGAGACAGTCCCTTAAAGCACCTCTATTGCT	3927		
Db	4008	CCAGGTTTCTTCTCATACGCTTCTTGAAGAGACAGTCCCTTAAAGCACCTCTATTGCT	3952		
Qy	3928	ACAAATTAAGTCTAGCAGATTTGTAACTTTAGTTTACAAAGTTCTAGAGATGCTTTTGGC	3987		

Db	3951	ACAAATTAAGTCTAGCAGATTTGTAACTTTAGTTTACAAAGTTCTAGAGATGCTTTTGGC	3892		
Qy	3988	TATGCTACACTGAAAGACTACAGCAACAAGAAATGATCCATCTTATCTCATCTCTGAA	4047		
Db	3891	TATGCTACACTGAAAGACTACAGCAACAAGAAATGATCCATCTTATCTCATCTCTGAA	3832		
Qy	4048	GCTTTGGCATCTCCAGCAAAAGATGTGCTATTTTACTGATACCATCACCATGAAGCCCAAC	4107		
Db	3831	GCTTTGGCATCTCCAGCAAAAGATGTGCTATTTTACTGATACCATCACCATGAAGCCCAAC	3772		
Qy	4108	AGTTTGGAGTCCAGATTAACCAAGCAGGTTTCAATGAAGCCCTTAAGTTATGATCATTA	4167		
Db	3771	AGTTTGGAGTCCAGATTAACCAAGCAGGTTTCAATGAAGCCCTTAAGTTATGATCATTA	3712		
Qy	4168	GATAAAGAGATTTATTTAGTCTCTATTAATCAAAATACCTCTGCAACGATCTTCTCAGTG	4227		
Db	3711	GATAAAGAGATTTATTTAGTCTCTATTAATCAAAATACCTCTGCAACGATCTTCTCAGTG	3652		
Qy	4228	CGGTCCATGGTCCAGTGCCACATATATGGGGTTCAGATGATTTACATTTGGTCTTGTCTCTC	4287		
Db	3651	CGGTCCATGGTCCAGTGCCACATATATGGGGTTCAGATGATTTACATTTGGTCTTGTCTCTC	3592		
Qy	4288	CCGGTGGATATAAATGATATTTCCAGGTAAAGATATTCCTTATTTTCAGACAAAAAAC	4347		
Db	3591	CCGGTGGATATAAATGATATTTCCAGGTAAAGATATTCCTTATTTTCAGACAAAAAAC	3532		
Qy	4348	ATACCAACCATCATGATCGAGGTGCAAGAGCATTTGCCCATGATGCAGAGGCTTTTCCA	4407		
Db	3531	ATACCAACCATCATGATCGAGGTGCAAGAGCATTTGCCCATGATGCAGAGGCTTTTCCA	3472		
Qy	4408	TCTGGAATCGAGGCTCTTGTAAAAAAATTTCTTTTCACTTGTCTGACAGACAGATGAGTCTT	4467		
Db	3471	TCTGGAATCGAGGCTCTTGTAAAAAAATTTCTTTTCACTTGTCTGACAGACAGATGAGTCTT	3412		
Qy	4468	ACGGAATAATGAAATTCATTCATTCAGATGCCTCTCTGTGTTTTTGAAGATACAGAAAGAC	4527		
Db	3411	ACGGAATAATGAAATTCATTCATTCAGATGCCTCTCTGTGTTTTTGAAGATACAGAAAGAC	3352		
Qy	4528	ACTGGACTACAGGAACATACAGATGATACTGCTTTTATTTGCTGTGTTTGAATTTCTG	4587		
Db	3351	ACTGGACTACAGGAACATACAGATGATACTGCTTTTATTTGCTGTGTTTGAATTTCTG	3292		
Qy	4588	GGTTTCCAGCCCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTCAAGATATT	4647		
Db	3291	GGTTTCCAGCCCAACCAACTGAGTGCAATATGTAGTCAATTCAGACTTTTCAAGATATT	3232		
Qy	4648	CCATATTCGATTTGGTGTGAGCAGACTATCCATAATCCTTTTGAAGTGGTTCCTCTTAAG	4707		
Db	3231	CCATATTCGATTTGGTGTGAGCAGACTATCCATAATCCTTTTGAAGTGGTTCCTCTTAAG	3172		
Qy	4708	TTTTTCGGGATTTCTGATCCAGTGATGGGTGCTCTCAAGAGGCTCAGCTAGCAGCACC	4767		
Db	3171	TTTTTCGGGATTTCTGATCCAGTGATGGGTGCTCTCAAGAGGCTCAGCTAGCAGCACC	3112		
Qy	4768	AAAAGCACAGAAATTTGTACTAGGTGTTAAAACAATTCAGATGATACACCAATGTGCCGT	4827		
Db	3111	AAAAGCACAGAAATTTGTACTAGGTGTTAAAACAATTCAGATGATACACCAATGTGCCGT	3052		
Qy	4828	ATACTCCTTCGCAAAAGTTCTAAGATAGTCAATTAATTTAGTGTAGTTTCAAGTTCAACT	4887		
Db	3051	ATACTCCTTCGCAAAAGTTCTAAGATAGTCAATTAATTTAGTGTAGTTTCAAGTTCAACT	2992		
Qy	4888	AAATGCTATGAGACTGGCTTTTAACAATTAAGGAGATGATCCTCAAAATTTTGTATGAC	4947		
Db	2991	AAATGCTATGAGACTGGCTTTTAACAATTAAGGAGATGATCCTCAAAATTTTGTATGAC	2932		
Qy	4948	ATATGCTTTTACTCTCAGGTTTCCCATTTGCTGTCACTGTGACATTCAGACTTCCGCTGT	5007		
Db	2931	ATATGCTTTTACTCTCAGGTTTCCCATTTGCTGTCACTGTGACATTCAGACTTCCGCTGT	2872		
Qy	5008	CGAGGTTTCATACAGAAATTTATTTCAAGATGTACAGTTTCTCAAAATGCAATGAAGACCA	5067		


```
Qy 4588 GGTTCAGCCAGCAACAACTGAGTCAATATGTAGTCAATCAGACTTTCAAGATATT 4647
Db 3237 GGTTCAGCCAGCAACAACTGAGTCAATATGTAGTCAATCAGACTTTCAAGATATT 3178
Qy 4648 CCATATCTGATTTGGTGTGACGACATATCCATATCTTTAGAGTGGTTCCTCTTAAG 4707
Db 3177 CCATATCTGATTTGGTGTGACGACATATCCATATCTTTAGAGTGGTTCCTCTTAAG 3118
Qy 4708 TTTTCGGGGAATTTCTGATGACGATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGACCC 4767
Db 3117 TTTTCGGGGAATTTCTGATGACGATGGGGTGTCTCAAGAAGGCTCAGCTAGCAGACCC 3058
Qy 4768 AAAAGCAGAAATTTGTTACTAGGTGTTAAACAATTCAGATGATACCAATGTGCCGT 4827
Db 3057 AAAAGCAGAAATTTGTTACTAGGTGTTAAACAATTCAGATGATACCAATGTGCCGT 2998
Qy 4828 ATACTCTTCCGAAGAGTTCTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTCAACT 4887
Db 2997 ATACTCTTCCGAAGAGTTCTAAGATTAGTCAATTAATTTGAGTAGTTCAGTTCAACT 2938
Qy 4888 AAATGTCTATGACACTGGGCTTTAAACAATTAAGAGAGATATCTCAAAACATTTGATGAC 4947
Db 2937 AAATGTCTATGACACTGGGCTTTAAACAATTAAGAGAGATATCTCAAAACATTTGATGAC 2878
Qy 4948 ATATGCCCTTTACTCTGAGGTTTCCCATTTGCTGTCACTGACATTCAGACTTCCGCTGT 5007
Db 2877 ATATGCCCTTTACTCTGAGGTTTCCCATTTGCTGTCACTGACATTCAGACTTCCGCTGT 2818
Qy 5008 CGGAGGTTTCATACAGAATTTATTTCAAGATGATCAGTTTCTAACAATGCATGAAGAAGCA 5067
Db 2817 CGGAGGTTTCATACAGAATTTATTTCAAGATGATCAGTTTCTAACAATGCATGAAGAAGCA 2758
Qy 5068 GAGGCTGTGTTGGCAACACCAACCAACCACTATAGTTGATATCATCTGCTGAATCTCTGA 5127
Db 2757 GAGGCTGTGTTGGCAACACCAACCAACCACTATAGTTGATATCATCTGCTGAATCTCTGA 2698
```

RESULT 5

```
US-10-357-930-26008/c
; Sequence 26008, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26008
; LENGTH: 4602
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: 1, 105, 4602
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-26008

Query Match 34.5%; Score 1768; DB 18; Length 4602;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 5; Indels 97; Gaps 1;

Qy 3148 GACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCCCTTGATATCAATGAAGATACAGAG 3207
Db 4580 GACCGGTTTGGCAGCAGCTCTACTAGTACATATTTTCCCTTGATATCAATGAAGATACAGAG 4521
Qy 3208 CCAACATTTTATGACCGATCTGGACCCATTAAGGATAAATAATTCATTTCCCTTTCTTTGCT 3267
Db 4520 CCAACATTTTATGACCGATCTGGACCCATTAAGGATAAATAATTCATTTCCCTTTCTTTGCT 4461
Qy 3268 TCTAGTAAATCTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCCCTAACAAAAACAT 3327
Db 4460 TCTAGTAAATCTGTGAAGAATCGTATCTTAAATTCGCTTACTTTGCCCTAACAAAAACAT 4401
Qy 3328 CGTAGTAGCAGTGCATCCAAAGAGGGAATTAATCATCTGAAAGTAAAGCAAGCAACAGG 3387
Db 4400 CGTAGTAGCAGTGCATCCAAAGAGGGAATTAATCATCTGAAAGTAAAGCAAGCAACAGG 4341
Qy 3388 CGAATCAGAACACCTTACGGAGCCAGTGTGATTTTAATCATAGTGAATTTTACACCC 3447
Db 4340 CGAATCAGAACACCTTACGGAGCCAGTGTGATTTTAATCATAGTGAATTTTACACCC 4281
Qy 3448 ATATCCACTGTA CAGAAAAACATTAATTAAGAGACTTTTATGGGAATAAGCAACT 3507
Db 4280 ATATCCACTGTA CAGAAAAACATTAATTAAGAGACTTTTATGGGAATAAGCAACT 4221
Qy 3508 GAGACACTGGTAGTACACCAAGCATTGGAGAAATGACTTAAATTCACCAAGAAATTTT 3567
Db 4220 GAGACACTGGTAGTACACCAAGCATTGGAGAAATGACTTAAATTCACCAAGAAATTTT 4161
Qy 3568 GGTACAGAGAATCAGAGAAAAATACAAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACG 3627
Db 4160 GGTACAGAGAATCAGAGAAAAATACAAGCCGAGAGAGGTTAGTAGTAGAAAGTTCAACG 4101
Qy 3628 AGCTCATATGAAGATACGTAGCCAAAAGTTTCAATACAGACACTACAACAAGTGGGATA 3687
Db 4100 AGCTCATATGAAGATACGTAGCCAAAAGTTTCAATACAGACACTACAACAAGTGGGATA 4041
Qy 3688 AGTTCAATGAGCTCAAGTCTTCCAGAGAGACAGTAGGTAGTATCAATCTATCGAC 3747
Db 4040 AGTTCAATGAGCTCAAGTCTTCCAGAGAGACAGTAGGTAGTATCAATCTATCGAC 3981
Qy 3748 ACAGACTGTGGAAGCATGAGTACTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTGCCT 3807
Db 3980 ACAGACTGTGGAAGCATGAGTACTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTGCCT 3955
Qy 3808 TTGACGCCACAGTCTAACCATCTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTGCCT 3867
Db 3954 TTGACGCCACAGTCTAACCATCTGTCTCTCTCAAAATCAAAATTCGGTGTCCCTGGTGCCT 3955
Qy 3868 CCAGGTTCTTCTCATACGCTTCTTAGAGAGACAGTCCCTTTAAAGCACCTCTATTGCT 3927
Db 3954 ---GGTTCTTCTCATACGCTTCTTAGAGAGACAGTCCCTTTAAAGCACCTCTATTGCT 3898
Qy 3928 ACAATTTAAAGTCTAGCAGATTTCTAACTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGC 3987
Db 3897 ACAATTTAAAGTCTAGCAGATTTCTAACTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGC 3838
Qy 3988 TATGCTACTGAAAAAGACTACAGCAACAAAGAAATGCATCCATCCTTATCTCACTCTGAA 4047
Db 3837 TATGCTACTGAAAAAGACTACAGCAACAAAGAAATGCATCCATCCTTATCTCACTCTGAA 3778
Qy 4048 GCTTTGGCATCTCCAGCAAAAGATGCTATTTTACTGATACCATCACCATGAAGGCCAAC 4107
Db 3777 GCTTTGGCATCTCCAGCAAAAGATGCTATTTTACTGATACCATCACCATGAAGGCCAAC 3718
```



```
Oy 4423 CTTGTAAAAAATCTTTT 4440
Db 552 GCTTTGNAAAAAATTTT 569

RESULT 7
US-09-814-353-6409
; Sequence 6409, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6409
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5, 368, 383, 423, 445, 485, 500, 538, 542, 549, 558
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-6409

Query Match          9.1%; Score 465.2; DB 10; Length 570;
Best Local Similarity 92.5%; Pred. No. 1.7e-115;
Matches 516; Conservative 0; Mismatches 38; Indels 4; Gaps 3;

Oy 3885 GCTTCCTAGAAAGACAGTCCCTTAAAGCACAGTCCCTTAAAGCACAGTCCCTTAAAGGTCTAGC 3944
Db 14 GCTTCCTAGAAAGACAGTCCCTTAAAGCACAGTCCCTTAAAGGTCTAGC 73

Oy 3945 AGATTGTAACCTTTAGTTACACAAGTCTTAGAGATGCTTTGGCTATGCTACACTGAAAG 4004
Db 74 AGATTGTAACCTTTAGTTACACAAGTCTTAGAGATGCTTTGGCTATGCTACACTGAAAG 133

Oy 4005 ACTACAGCAACAAAGAAATGCAATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGC 4064
Db 134 ACTACAGCAACAAAGAAATGCAATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGC 193

Oy 4065 AAAAGATGTGCTATTTACTGTATACATCAATCAATCAAGGCCAACAGTTTGGTCCAGATT 4124
Db 194 AAAAGATGTGCTATTTACTGTATACATCAATCAATCAAGGCCAACAGTTTGGTCCAGATT 253

Oy 4125 AACACCAAGCAGGTTTCATGAAGCCTTAAGTTATGTCATCATAGATAAGAGATTATT 4184
Db 254 AACACCAAGCAGGTTTCATGAAGCCTTAAGTTATGTCATCATAGATAAGAGATTATT 313

Oy 4185 GAGTCCTATTAAATCAAAATACCTTGCAA-CGATCTTCTCTCAGTCCGCTCCATGTGTCCTCA 4243
Db 314 GAGTCCTATTAAATCAAAATACCTTGCAAACCGATCTTCTCTCAGTCCGCTCCATGTGTCCTCA 373

Oy 4244 GTGCCACATATGGGGGTTTCAG-ATGATTACATTTGGTCTTGTCTCTCCCGGTGGATATAAAT 4302
```

```
Db 374 GTGCCACATNGGGGGGTCAAGATGATTACATTGGGCTTGCTCTCCCGNGGATATAAAT 433
Oy 4303 GATATATTTCCAGGTAAAGGATATTCCTTATTTTCCAGACAAAACATACACCATCATGAT 4362
Db 434 GATATATTTCCAGGTAAAGGATATTTCTTATTTTCCAGACAAAACAT-CCCCCATCATGAT 491
Oy 4363 GATCAGGTGCAAGAGCATTTGCCCATGATGACAGGAGTCTTCCATCTGGAACTGGAGGT 4422
Db 492 GATCCAGNGGCCAGACATTTGCCCTGAATGACAGGAGGCTTTTCATNTGNACTGGGNGG 551
Oy 4423 CTTGTAAAAAATCTTTT 4440
Db 552 GCTTTGNAAAAAATTTT 569

RESULT 8
US-10-357-930-9378/c
; Sequence 9378, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9378
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-9378

Query Match          7.5%; Score 385.6; DB 18; Length 452;
Best Local Similarity 99.0%; Pred. No. 7.2e-94;
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 4129 CCAAGCAGGTTTCATGAAGCCTTAAAGTTATGTCATCATTAGATAAAGAGATTTATTGAGT 4188
Db 400 CTTTCAGGTTTCATGAAGCCTTAAAGTTATGTCATCATTAGATAAAGAGATTTATTGAGT 341

Oy 4189 CCTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCTAGTGTCCAGTGCC 4248
Db 340 CCTATTAAATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCTAGTGTCCAGTGCC 281

Oy 4249 ACATATGGGGGTTTCAGATGATTACATTTGTCCTTCTCCGGTGGATATAAATGATATA 4308
Db 280 ACATATGGGGGTTTCAGATGATTACATTTGTCCTTCTCCGGTGGATATAAATGATATA 221

Oy 4309 TTCAGGTTAAAGGATATTCCTTATTTTCCAGACAAAACATACACCATCATGATGATCGA 4368
Db 220 TTCAGGTTAAAGGATATTCCTTATTTTCCAGACAAAACATACACCATCATGATGATCGA 161

Oy 4369 GGTGCAAGAGCATTTTGCCCATGATGACAGGAGGTCTTCCATCTGGAACCTGGAGGTCTTGTA 4428
```

Db 160 GGTGCAAGAGCAATTTCCCATGATGACGAGAGTCTTCCATCTGGAGCTCTTGTA 101
QY 4429 AAAAAATCTTTTCACTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATCAATC 4488
Db 100 AAAAAATCTTTTCACTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATCAATC 41
QY 4489 CATTGAGTGCCTCTCTGTTTTAGAAAGTAC 4520
Db 40 CATTCAATGCGCTCTCTGTTTTAGAAAGTAC 9

RESULT 9
US-10-357-930-30581/c
; Sequence 30581, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30581
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-30581

Query Match 7.5%; Score 383.4; DB 18; Length 427;
Best Local Similarity 99.7%; Pred. No. 2.7e-93;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4136 GGTTCATGAAAGCCTTAAGTTATGTCATCATTAGATAAAGAGATTTATTGAGTCTTATTA 4195
Db 427 GGTTCATGAAAGCCTTAAGTTATGTCATCATTAGATAAAGAGATTTATTGAGTCTTATTA 368

QY 4196 ATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCATGTCGAGTCCAGTGCCACATATG 4255
Db 367 ATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCATGTCGAGTCCAGTGCCACATATG 308

QY 4256 GGGGTTTCAGATGATTACATTTGGTCTTCTCCCGGTGGATATAAATGATATATTCAGG 4315
Db 307 GGGGTTTCAGATGATTACATTTGGTCTTCTCCCGGTGGATATAAATGATATATTCAGG 248

QY 4316 TAAAGGATATTCCTTATTTTCAGCAAAAAACATACACCATGATGATCGAGGTGCAA 4375
Db 247 TAAAGGATATTCCTTATTTTCAGCAAAAAACATACACCATGATGATCGAGGTGCAA 188

QY 4376 GAGCATTTGCCATGATGACGAGGTCTTCCATCTGGAACCTGGAGTCTTGTAAAAAAT 4435
Db 187 GAGCATTTGCCATGATGACGAGGTCTTCCATCTGGAACCTGGAGTCTTGTAAAAAAT 128

QY 4436 CTTTTCACCTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATTCATTCAG 4495
Db 127 CTTTTCACCTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATTCATTCAG 68

QY 4496 CTTTTCACCTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATTCATTCAG 4495

Db 127 CTTTTCACCTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATTCATTCAG 68
QY 4496 ATGCTCTCTGTTTTAGAAAGTAC 4520
Db 67 ATGCTCTCTGTTTTAGAAAGTAC 43

RESULT 10
US-10-357-930-39551/c
; Sequence 39551, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39551
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-39551

Query Match 7.5%; Score 383.4; DB 18; Length 427;
Best Local Similarity 99.7%; Pred. No. 2.7e-93;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4136 GGTTCATGAAAGCCTTAAGTTATGTCATCATTAGATAAAGAGATTTATTGAGTCTTATTA 4195
Db 427 GGTTCATGAAAGCCTTAAGTTATGTCATCATTAGATAAAGAGATTTATTGAGTCTTATTA 368

QY 4196 ATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCATGTCGAGTCCAGTGCCACATATG 4255
Db 367 ATCAAAATACCCCTGCAACGATCTTCTCAGTGGGTCCATGTCGAGTCCAGTGCCACATATG 308

QY 4256 GGGGTTTCAGATGATTACATTTGGTCTTCTCCCGGTGGATATAAATGATATATTCAGG 4315
Db 307 GGGGTTTCAGATGATTACATTTGGTCTTCTCCCGGTGGATATAAATGATATATTCAGG 248

QY 4316 TAAAGGATATTCCTTATTTTCAGCAAAAAACATACACCATGATGATCGAGGTGCAA 4375
Db 247 TAAAGGATATTCCTTATTTTCAGCAAAAAACATACACCATGATGATCGAGGTGCAA 188

QY 4376 GAGCATTTGCCATGATGACGAGGTCTTCCATCTGGAACCTGGAGTCTTGTAAAAAAT 4435
Db 187 GAGCATTTGCCATGATGACGAGGTCTTCCATCTGGAACCTGGAGTCTTGTAAAAAAT 128

QY 4436 CTTTTCACCTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATTCATTCAG 4495
Db 127 CTTTTCACCTTGTACGACAGCAGATGAGTCTTACGGAATAATGAATTCATTCAG 68

QY 4496 ATGCTCTCTGTTTTAGAAAGTAC 4520

Db 67 ATGCCTCTCTGTTTTAGAAAGTAC 43
|||||

RESULT 11
US-10-930-930-209/c
; Sequence 209, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382_feature
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-209

Query Match 6.6%; Score 336.4; DB 18; Length 416;
Best Local Similarity 98.7%; Pred. No. 1.7e-80;
Matches 370; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
Qy 4149 CTTAAGTTATGCATCATTTAGATAAAGAGATTANTGAGTCCTATTAAATCAAAATACCC 357
Db 416 CTTAAGTTATGCATCATTTAGATAAAGAGATTANTGAGTCCTATTAAATCAAAATACCC 357
Qy 4208 TGCAACCATCTT-CCTCAGTGGGTCCATGTGTCCAGTGCACATATGGGGTTTCAGAT 4266
Db 356 TGCAACCATCTT-CCTCAGTGGGTCCATGTGTCCAGTGCACATATGGGGTTTCAGAT 297
Qy 4267 GATTACATTTGCTCTTCTCCCGTGGATATAATGATATATTTCCAGGTAAAGGATATT 4326
Db 296 GATTACATTTGCTCTTCTCCCGTGGATATAATGATATATTTCCAGGTAAAGGATATT 237
Qy 4327 CCCTATTTTCAGACAAAACATACCAACCATGATGATCGAGTGCAGAGCATTTTGGC 4386
Db 236 CCCTATTTTCAGACAAAACATACCAACCATGATGATCGAGTGCAGAGCATTTTGGC 177
Qy 4387 CATGATCGAGGAGTCTTCCATCTGGAACCTGGAGGCTTGTAAAAATTTCTTTTCATTG 4446
Db 176 CATGATCGAGGAGTCTTCCATCTGGAACCTGGAGGCTTGTAAAAATTTCTTTTCATTG 117
Qy 4447 CTACGACGACGATGAGTCTTACCGAAATATGAAATTCATTCAGATGCCTCTC-T 4505
Db 116 CTACGACGACGATGAGTCTTACCGAAATATGAAATTCATTCAGATGCCTCTC-T 57
Qy 4506 GTTTTAAAGTAC 4520

Db 56 GTTTTAAAGTAC 42
|||||

RESULT 12
US-09-814-353-12794
; Sequence 12794, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12794
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-12794

Query Match 6.5%; Score 334; DB 10; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.9e-80;
Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3885 GCTTCTTAGAAGACACAGTCCCTTAAAGCACCCTCTATTGCTACAATTTAAAGTCTAGC 3944
Db 8 GCTTCTTAGAAGACACAGTCCCTTAAAGCACCCTCTATTGCTACAATTTAAAGTCTAGC 67
Qy 3945 AGATTGTAACCTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGCTATGCTACACTGAAAAG 4004
Db 68 AGATTGTAACCTTTAGTTTACACAAGTCTTAGAGATGCTTTTGGCTATGCTACACTGAAAAG 127
Qy 4005 ACTACAGCAACAAGAAATGCATCCATCCTTTATCTCAGTCTGAAAGCTTTGGCATCTCCAGC 4064
Db 128 ACTACAGCAACAAGAAATGCATCCATCCTTTATCTCAGTCTGAAAGCTTTGGCATCTCCAGC 187
Qy 4065 AAAAGATGCTCTATTCTTACTGATACCATCACCATGAAAGCCCAACAGTTTGGAGTCCAGATT 4124
Db 188 AAAAGATGCTCTATTCTTACTGATACCATCACCATGAAAGCCCAACAGTTTGGAGTCCAGATT 247
Qy 4125 AACACCAAGAGGTTCTATGAAAGCCTTAAGTTATGATCATTTAGATAAAGAGATTATT 4184
Db 248 AACACCAAGAGGTTCTATGAAAGCCTTAAGTTATGATCATTTAGATAAAGAGATTATT 307
Qy 4185 GAGTCTCTATTAAATCAAAATACCTGCAACGATCT 4218
Db 308 GAGTCTCTATTAAATCAAAATACCTGCAACGATCT 341

RESULT 13
US-09-933-797-314/c
; Sequence 314, Application US/09933797
; Patent No. US20020155119A1
; GENERAL INFORMATION:
; APPLICANT: Robert A. Sikes et al.


```
; TITLE OF INVENTION: Isolation and Use of Fetal Urogenital
; FILE REFERENCE: 9901-007-999
; CURRENT APPLICATION NUMBER: US/09/933,797
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US/09/482,933
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: PCT/US99/10746
; PRIOR FILING DATE: 1999-05/14
; PRIOR APPLICATION NUMBER: 60/085,383
; PRIOR FILING DATE: 1998-05-14
; NUMBER OF SEQ ID NOS: 811
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 314
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Murine
US-09-933-797-314

Query Match          6.3%; Score 321; DB 9; Length 450;
Best Local Similarity 83.4%; Pred. No. 2.9e-76;
Matches 373; Conservative 4; Mismatches 67; Indels 3; Gaps 1;

QY 2929 AAACCAACAAGGCTGTGATATTTAAATGTCAACAACCTGGGATGCTGTGGGCGATGT 2988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 AAACCAAGGTAAGGCTGTGATATTTGAAATGTCAACAGCTGGGATTTCTGTAAAGS---ACA 394

QY 2989 CGCAACATCTGTGGCCAGTGGTTCAGATGATGTGGAACAACCTCTGTAATGAACTTTCA 3048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 GTGDADGDTCTATGGCCAGTGGTTCAGATGATGTGGGAGCGCTCTGCAATGAGCTCTCA 334

QY 3049 TCTATCCCAAGCACTCTAAATTTGAATCGGAGTCAACAGCTCTAGACATATAGTGA 3108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 333 TCTGTCCCAAGCACTCTAGCTTGAATCTGAGTCACCACTCTTAGACATATACAGTGAG 274

QY 3109 AGTGAATCTGTGCATCGATATGTTCAATATGGAGGATGACCGGTTTGGCAGCAGCTCT 3168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 AGTGAGCTGTCACCATCAAGCATGTTTCATGCTGGAGGAGCAGCGCTTTCGGCAGCCTCC 214

QY 3169 ACTAGTACATTTTCTTCATATCAATGAAGATGATGACAGCAACATTTTATGACCGATCT 3228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 ACCAGACATATTTTCTTCATATCAATGAAGATGCTGGAGCAGCATTTTATGACCGACCT 154

QY 3229 GGACCCATAAAGGATAAAATTCATTCCTTTCTTTGTTCTAGTAAATCTTGTGAAGAT 3288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 GGACCCATAAAGATATAAATTCATTTCTTTTGGCTCCAGCBAACCTTGTAAAGAT 94

QY 3289 CGTATCTTAAATTCGCTTACTTTGCTTAAACAAAACATCGTAGTAGCAGTATCCAAAA 3348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 CGTATCTTAAATTCATTTACTTTGCTTAAACAAAACCGCAGTAGCAGTATCCAAAA 34

QY 3349 GGAGGGAATATCATCTGAAGTAG 3375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 GGAGGGAATATCTGATCTGAAAATAAG 7
```

```
RESULT 14
US-09-910-943-73
; Sequence 73, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/1G1480S1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Xenopus laevis
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(726)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-73

Query Match          5.5%; Score 280.2; DB 9; Length 726;
Best Local Similarity 69.5%; Pred. No. 5.2e-65;
Matches 459; Conservative 0; Mismatches 179; Indels 22; Gaps 5;

QY 4342 AAAAAACATACCAACCATGATGATGAGGTGCAAGAGCAATTTGCCCATGATGAGGAGGT 4401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 75 AAAGCGGTGCTCCCAAGTGAAGAGGAGCAGCTAAAATGTTTGCACCTGAAGTAGGAGGC 134

QY 4402 CTTCCATCTGGAACCTGGAGGTCTTTTAAAAAATTTCTTTTCACTTGTGACAGCAGATG 4461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 ATCCAA-----GGATCCGGGGTGAAGAGTCCGTACACGGCCTTGGCTCAGCAGTTTC 185

QY 4462 AGTCTTACGGAATATGAATTCATATCCATTCAGATGCTCTCTGTTTTTAGAAAGTACA 4521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 AGCATTTACGAGATATATGAATCTCGATGCTGTCGGATGCTCAGATTTCTAGAGATACA 245

QY 4522 GAAGACATCTGGACTACAGGAACATACAGATGATACTGCTTTTATTTGCTGTGATGAA 4581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GAAGACACGGGCTGCGAGGAGCACACTGATGATACTGCTGTCTGTGTTGGCATTCAC 305

QY 4582 ATTCTGGGTTTCCAGCCAGCAACCACTGAGTGCATATGTAGTCAATTCAGACTTTCAA 4641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 TTAATGGGTTACAGCCAGTCAAAATCAGTTTCAATG---AATATAGCCGTTTAGACTTTCC 362

QY 4642 GATATTTCCATATTTCTGATTTGTTGAGCAGACTATCCATAATCTTTAGAAGTGGT--- 4697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 GACATTTCCATATTTCTGATGCTGGTGGCTCATATATCCCAATCACTTAGAGGTTGTTATCC 422

QY 4698 --TCCTCTTAAGTTTTCGGGATTTCTGGATGCAGTATGTTGGGTGCTCAAGAGGCTCA 4755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 CATTTCTTCAAGTTTCTGGGATATCAGGTTGATGATGTTGTTGTTCC---AGGCTCG 479

QY 4756 GCTAGCAGCACCAAGACAGAAATTTGTTACTAGTGTATAAACAATTCACAGATGATACA 4815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCAAGCAGTGAACAAGATACAGATGTTAGTTCAGGTGGCAAAATCAATTCGCAAGATACA 539

QY 4816 CCAATGTGCGTATATCTCTCGCAAGAAAGTTCTAAGATTAAGTCAATTAATTTGAGTAGT 4875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 CTGTTTGCAGATATTAATCTCCGGAAGAGGCTCTTAAGACTTGTAACTTAACCTGAGTAGC 599

QY 4876 TCAGTTTCACTAAATGTATGAGACTGGGCTTTTAAACAATTAAGAGAGATATCTCAA 4935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 TCCGTAGGAACGAAAGGCCATGAAACTGGGCTCTTAAACGATTAAGGAGAGTTTCTCAA 659

QY 4936 ACATT-TGATGACATATGCTTTTACTCTGAGGTTTCCCATTTGCTGTCACTGCACATT 4994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 GCCTTTTGATGACATTTGCTCTTATTTCTGAGGTTTCCCACTTTATTAGCACATTGCAATT 719
```

```
RESULT 15
US-10-029-386-8811
; Sequence 8811, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8811
; LENGTH: 567
; TYPE: DNA
```

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC008964.4

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59

; OTHER INFORMATION: SWISSPROT HIT: P40061, EVALUE 3.20e+00

; OTHER INFORMATION: NT HIT: g115828471, EVALUE 5.40e-01

; OTHER INFORMATION: EST_HUMAN HIT: B1601569.1, EVALUE 3.00e-71

US-10-029-386-8811

Query Match	3.3%;	Score 166.8;	DB 16;	Length 567;
Best Local Similarity	93.5%;	Pred. No. 3.6e-34;		
Matches 174;	Conservative 0;	Mismatches 12;	Indels 0;	Gaps 0;
Qy	568	ATTATCTGTGAACACTTCAGACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAAC	627	
Db	260	AGTTTTCCTCTTTTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAAC	319	
Qy	628	ACCATATTGAAAAATGTGATTGATTCGCAATTAAGTCGAATAAATGAGGCCCTTAATTACT	687	
Db	320	ACCATCTTGAATAATGTGATTCGCAATTAAGTCGAATAAATGAGGCCCTTAATTACT	379	
Qy	688	ACAAATTTGCAACCTCTTAATCATCCAAAGACTCGGCAGTATGTCCGAGCTGATGTAGAA	747	
Db	380	ACAAATTTGCAACCTCTTAATCATCCAAAGACTCGGCAGTATGTCCGAGCTGATGTAGAA	439	
Qy	748	TTAGAG 753		
Db	440	TTAGAG 445		

Search completed: April 9, 2005, 15:10:29
Job time : 1854.73 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 17:12:23 ; Search time 10486.4 Seconds
(without alignments)
18610.264 Million cell updates/sec

Title: US-10-782-244-2
Perfect score: 5127
Sequence: 1 atggcgcgatcgccgcgg.....atacatctgtgaatcctga 5127

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4860.6	94.8	7311	3	CR749280 Homo sapi
2	2254	44.0	4696	3	AL834497 Homo sapi
3	2123.6	41.4	2497	3	AK036149 Mus muscu
4	2116.8	41.3	3881	3	AK082586 Mus muscu
5	1047.6	20.4	1475	9	AY4117033 Homo sapi
6	880.8	17.2	1453	5	AY4117035 Mus muscu
7	824.8	16.1	872	5	BQ049342 AGENCOURT
8	818.6	16.0	1977	3	AK075662 Mus muscu
9	806.4	15.7	1211	3	AK087858 Mus muscu
10	753.2	14.7	878	6	CD653846 AGENCOURT
11	732.2	14.3	856	4	BG623200 602648109
12	710.2	13.9	867	4	BG414183 GKP2-237
13	702.6	13.7	893	5	BU746120 CH3#002 A
14	701.8	13.7	831	4	BI601569 CH3#038 F
15	693.8	13.5	1377	9	AY4117034 Pan trogl
16	690.8	13.5	801	7	CF723187 UI-M-GV0-
17	681	13.3	3571	3	AK019537 Mus muscu
18	681	13.3	4697	3	AK029196 Mus muscu
19	680	13.3	807	5	BU751233 CH3#038 F
20	679.4	13.3	1632	3	AK054119 Mus muscu
21	640.8	12.5	645	7	CV024787 2284 Full
22	638.4	12.5	822	5	BU746121 CH3#002 A
23	636.6	12.4	747	7	CR629809 DKFP2469N
24	627.8	12.2	807	5	BU751232 CH3#038 F

25	625	12.2	825	6	CD657315	CD657315 AGENCOURT
c	26	617.2	12.0	674	5	EX6667705 BX6667705
c	27	612.6	11.9	659	5	EX666671 BX666671
	28	609.8	11.9	676	4	BI463004 BI463004
	29	608.6	11.9	748	7	CN534331 UI-M-H00-
	30	599.8	11.7	791	7	CK637702 UI-M-H00-
	31	587	11.4	835	6	CD655466 AGENCOURT
	32	584	11.4	700	1	AU253730 AU253730
	33	579	11.3	671	6	CB434014 610595 MA
	34	571.2	11.1	841	7	CN535250 UI-M-H50-
	35	566.6	11.1	664	2	BB613723 BB613723
	36	562.4	11.0	651	6	BY722545 BY722545
	37	561	10.9	581	5	BP314047 BP314047
	38	560.4	10.9	581	5	BP352173 BP352173
	39	560.4	10.9	582	5	BP314060 BP314060
	40	560	10.9	582	5	BP352842 BP352842
	41	558	10.9	581	5	BP199668 BP199668
	42	541.8	10.6	706	1	AV752256 AV752256
	43	541.6	10.6	640	6	BY740237 BY740237
	44	537.6	10.5	581	5	BP195123 BP195123
	45	536.2	10.5	734	7	CF743854 UI-M-G10-

ALIGNMENTS

RESULT 1
CR749280 7311 bp mRNA linear HTC 19-AUG-2004
LOCUS Homo sapiens mRNA; cDNA DKFP686B11164 (from clone DKFP686B11164).
DEFINITION CR749280
ACCESSION CR749280.1 GI:51476289
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 7311)
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo G., Han, M. and Wiemann, S.
CONSRM The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the CDNA
sequencing consortium of the German Genome Project. This clone
(DKFP686B11164) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFP686B11164
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..7311
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686B11164"
/tissue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="hypothetical protein, differentially spliced"
1..7311
/gene="DKFP686B11164"
866..5023
/gene="DKFP686B11164"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18135.1"

/db_xref="GI:51476290"
/translation="MGIATPRSWAGIINLCKPGNSIQSLIGVLCPINMEIRRLGLE
VLYDIPPLPVPVTEFEALISVDPGRFQDSWRSLSDGFVAEAKTILPHRARSPL
MNYLALLISAFIRNGLLEGLVEVITNSDDHISVRATILLEGILHMAINTILPHSHSH
LHCLPTLMMASFOIPKEKRLRASAALNCKRHEMKRGPYSLHLHIIQKATA
THKRDQYLRVQKDFILKDHTEALLNLRDSQVLKHENKRWGNWNLITLLEKFNVN
LNYKDEQLHRFVRLLLYFYKPSKLYANLRLDLFAKQLTAVGQFTFELISEEDG
KQYDLVKDIWOLNASSGMKPERSLONGLLATLSOYFELFTGLSHPHVKMLE
KCSVQCLNLCKLNQDHLKLTVSSLDYSRDLGLARVLSKILTAATDACLRYATKH
LRVLKRVNVEFNWGIIBELLTQLDHKNKTISSREALDILDEACEDKANLHALIOMKPA
LSHLGDKLLLRLLSLPSKGLSYLNERGYVAKQLEKWHREYNPKYVDLLEQNEAL
TYRREPVDGNVRRNSRLQPHVILPHLYQLVHHKTHCHLELVONILTELCRVN
RTPDLKWEIKKALWALGNISWNLQJENVIPIDLILKAKQCEVLSIRGT
CVVLGLAKTGQDILKCHNDVAVRSHKHLNVPVDDVEQDLCNELSSPSTLSIN
SSTSRHNSSESVPMSFLEDDRFGSSSTSPFLDINEDTEPTFYDRSGPIKDNK
SPPFPASSKLKVNRLINSILTPNKHRSDDPKGKLSSEKTSNRRIRLTLEPSVDF
NHSDDPFTSTVQKILQLETPMGNKHIEDTGSTPSIGENDLKPFTNPGTENHRENTS
RELVVESTSSHMKIRSGSFNTDTTIGSISMSSPRETVGDVATMTMDDCGSMST
VVSSTKITSHYLTQSNHLSLSKSSVSLVPPGSSHTLPPRAQSLKAPSATIKSLA
DCNFSYTSRDAFGVATLRLQOQRMHPSLSHSEALASPAKDVLPDTITLTKANSFES
RLTPSRFKALSVASLDKEDLLSPINQNTLQSSSVRSMVSSATYGGDDYIGLALPV
DINDLFOVKDIPIVOTKNTPPHDDRGARAFADHAGESTEDTGLQEHDTDNCLYCVIE
ILGQPSNOLSAICHSDFODJ.PYSDMCEQTLHNPVVPKFSGIGSCDGVSOEGS
ASTSTEGLLGKVIIPDDTPMCRILLRKEVLRLVINLSSWSKTEHGLLTKEKY
POTFDICILYSEVSHLLSHCTPLRPPCRFIQELFDQVDFLOWHEBAEAVLATPPKQPI
VDTSAES"

ORIGIN

Query Match	94.8%;	Score	4860.6;	DB	3;	Length	7311;
Best Local Similarity	97.6%;	Pred. No.	0;				
Matches	5002;	Conservative	0;	Mismatches	9;	Indels	116;
Gaps	2;						
QY	1	ATGGCGCGCATCGCGCGCGCGCTCTCTGAAGAACTCCGAGTACGAGGGCGGAATGAC	60				
Db	13	ATGGCGCGCATCGCGCGCGCGCTCTCTGAAGAACTCCGAGTACGAGGGCGGAATGAC	72				
QY	61	AGCGCGAGGAGAACGTCCCGCTGGATCTGACCCGAGAACCTTCTGATAAATTTAGAGAG	120				
Db	73	AGCGCGAGGAGAACGTCCCGCTGGATCTGACCCGAGAACCTTCTGATAAATTTAGAGAG	132				
QY	121	ATTCCTCAAAATGTGGCCAGATGTCAGGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG	180				
Db	133	ATTCCTCAAAATGTGGCCAGATGTCAGGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG	192				
QY	181	AATAAATCTTTACTAAGCTTCTTTGTGATATTTGGCCACAGTGAAGAAAACTGGGCTTTTAC	240				
Db	193	AATAAATCTTTACTAAGCTTCTTTGTGATATTTGGCCACAGTGAAGAAAACTGGGCTTTTAC	252				
QY	241	TATGAGGATATCAAAATTTGTTTCGGTGTAGCTTTTAAATCAAGCAAAAGAAAGTGGGA	300				
Db	253	TATGAGGATATCAAAATTTGTTTCGGTGTAGCTTTTAAATGAGCAAAAGAAAGTGGGA	312				
QY	301	GCAGCAGGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGTATTTCTCCAGAAAGTG	360				
Db	313	GCAGCAGGGCTACGAGCGCTTCGATATCTCATCAAGACTCCAGTATTTCTCCAGAAAGTG	372				
QY	361	CTAAAAATGAAAGTGGACTATTTAATAGCTAGGTGCAATGACATACAAACAGAGCAACGAG	420				
Db	373	CTAAAAATGAAAGTGGACTATTTAATAGCTAGGTGCAATGACATACAAACAGAGCAACGAG	432				
QY	421	GTAGAGAGGACACAGCACTTCGATTTAGTCAAGAAAGATGATTTACTGTGAATGCTTCCTTG	480				
Db	433	GTAGAGAGGACACAGCACTTCGATTTAGTCAAGAAAGATGATTTACTGTGAATGCTTCCTTG	492				
QY	481	TTTCTTAGTTCGTGACCAACTCATTAATTTGAGTTGGAATGATGACTTTCAGAAAGA	540				
Db	493	TTTCTTAGTTCGTGACCAACTCATTAATTTGAGTTGGAATGATGACTTTCAGAAAGA	552				
QY	541	GACAGAAATGGTCCGAGAGTGCATTATCTGTGAACCTTAGCACTTTCAGAAATCCAGAG	600				
Db	553	GACAGAAATGGTCCGAGAGTGCATTATCTGTGAACCTTAGCACTTTCAGAAATCCAGAG	612				
QY	601	GTGGTGGCCCTTCGAGGAGGAGCTAAACACCATATTTGAAAAAATGTGATTTGCCAATTA	660				

Db	613	GTGGTGGCCCTTCGAGTGGACTAAACACCACTTTGAAAAATGTGATTCGATTCGCAATTA	672
Qy	661	AGTCGAATAAATGAGGCCCTTAATTAATAAATTTTGCACTTCTTAATCATCCAAAGACT	720
Db	673	AGTCGAATAAATGAGGCCCTTAATTAATAAATTTTGCACTTCTTAATCATCCAAAGACT	732
Qy	721	CGGCGATATGTGGAGCTGATAGTAATTAGAGAGAAATTTTAGCACCTTATCTGATTTT	780
Db	733	CGGCGATATGTGGAGCTGATAGTAATTTAGAGAAATTTAGCACCTTATCTGATTTT	790
Qy	781	CACCTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAGAGACAGAGAGAACGACGA	840
Db	791	CACCTACAGACATAGTCCAGATACAGCTGAAGGACAGCTCAAGAGACAGAGAGAACGACGG	850
Qy	841	TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGGATCATGGGAGGTATTTAAT	900
Db	851	TTTCTAGCCAGTAAATGGGAATCATAGCAACATTCGGATCATGGGAGGTATTTAAT	910
Qy	901	TTATGTAAACCTCGAAATTCCTGGATCCAGTCTCTAATAGGAGTACTTTGCATACCAAAAT	960
Db	911	TTATGTAAACCTCGAAATTCCTGGATCCAGTCTCTAATAGGAGTACTTTGCATACCAAAAT	970
Qy	961	ATGGAATAAAGCGGAGCTCTACTGAAAGTCTTTATGATATATTTCTCTCTACCT	1020
Db	971	ATGGAATAAAGCGGAGCTCTACTGAAAGTCTTTATGATATATTTCTCTCTACCT	1030
Qy	1021	GTTCGTGATGAGGAGTTCATAGAGCACTACTCACTGCTAGATCCAGGAGGTTCCAAGAC	1080
Db	1031	GTTCGTGATGAGGAGTTCATAGAGCACTACTCACTGCTAGATCCAGGAGGTTCCAAGAC	1090
Qy	1081	AGTTGGAGGCTTTTCCAGATGGCTTTTGTGGAGCTGAGGCAAAACATTTCTCTCATCGT	1140
Db	1091	AGTTGGAGGCTTTTCCAGATGGCTTTTGTGGAGCTGAGGCAAAACATTTCTCTCATCGT	1150
Qy	1141	GCCAGATCCAGGCGACCTCATCGATAATTTATTTGGCACTGATCTCTGCAATTTAT	1200
Db	1151	GCCAGATCCAGGCGACCTCATCGATAATTTATTTGGCACTGATCTCTGCAATTTAT	1210
Qy	1201	CGTAATGACATTTTAGAGGGTCTAGTTGAAGTGTATAACAAACAGTGTATCATCTCA	1260
Db	1211	CGTAATGACATTTTAGAGGGTCTAGTTGAAGTGTATAACAAACAGTGTATCATCTCA	1270
Qy	1261	GTTAGAGCTACCATCTCTTTTAGGAGAGCTTTTACATATGGCAAAACATTTCTCTCAT	1320
Db	1271	GTTAGAGCTACCATCTCTTTTAGGAGAGCTTTTACATATGGCAAAACATTTCTCTCAT	1330
Qy	1321	TCACATAGCCATCATTTACACTGCTTGCAACCTTAATGAATATGGCTGCAATCTCTTGCAT	1380
Db	1331	TCACATAGCCATCATTTACACTGCTTGCAACCTTAATGAATATGGCTGCAATCTCTTGCAT	1390
Qy	1381	ATCCCCAAGGAAAAGAGACTGCGAGCCAGTGCAGCTTTGAACTGTTTAAAAACGCTTCCAT	1440
Db	1391	ATCCCCAAGGAAAAGAGACTGCGAGCCAGTGCAGCTTTGAACTGTTTAAAAACGCTTCCAT	1450
Qy	1441	GAAATGAAGAAACGAGGACCTTAAGCTTTATAGTCTTCAATTTAGACCACTATTTCAGAAA	1500
Db	1451	GAAATGAAGAAACGAGGACCTTAAGCTTTATAGTCTTCAATTTAGACCACTATTTCAGAAA	1510
Qy	1501	GCAATTTGCAACACACAGAAACGGGATCAGTATCTCCGAGTTTCAGAAAGATATATTATC	1560
Db	1511	GCAATTTGCAACACACAGAAACGGGATCAGTATCTCCGAGTTTCAGAAAGATATATTATC	1570
Qy	1561	CTTAAAGGATACAGAGAGCTCTTTTAAATTAACCTTAGAGATAGCAAGTCTCTCAACAT	1620
Db	1571	CTTAAAGGATACAGAGAGCTCTTTTAAATTAACCTTAGAGATAGCAAGTCTCTCAACAT	1630
Qy	1621	AAAGCAATCTTGAATGGAATCTTATAGGAGCCATTTCTTAAGTGGGCAAAATGTA	1680
Db	1631	AAAGCAATCTTGAATGGAATCTTATAGGAGCCATTTCTTAAGTGGGCAAAATGTA	1690
Qy	1681	AATCTAAGAAACTATTAAGATGAACAGTTTACAGGTTTGTAGGAAGACTACTTTATTTT	1740

Db 1691 ATCTAAGAACTATAAGATGAAACAGTTACACAGTTTGTACGAAGACTACTTTATTTT 1750
Qy 1741 TACAAGCCAGCAGTAATAATTATATGCCAAACCTGGATCTGGATTTGCCAAGGCCAAACAG 1800
Db 1751 TACAAGCCAGCAGTAATAATTATATGCCAAACCTGGATCTGGATTTGCCAAGGCCAAACAG 1810
Qy 1801 CTCAGGTTGTAGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGGCA 1860
Db 1811 CTCAGGTTGTAGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAAGAGGATGGGCA 1870
Qy 1861 GCTACTTATAGAGATCTAGTAAGGATATTGTTTCACTGCTCAATGCTTCACTCTGGAATG 1920
Db 1871 GCTACTTATAGAGATCTAGTAAGGATATTGTTTCACTGCTCAATGCTTCACTCTGGAATG 1930
Qy 1921 AAACCCGAAAGAGTCTTCAAAATTAATGTTTATGACACCCCTTTAGTCAACACTACTTT 1980
Db 1931 AAACCCGAAAGAGTCTTCAAAATTAATGTTTATGACACCCCTTTAGTCAACACTACTTT 1990
Qy 1981 TTATTTATGGNACACTTTCTTGGCCACCTCATGAGGTTAAATGCTGGAATAATGCAAT 2040
Db 1991 TTATTTATGGNACACTTTCTTGGCCACCTCATGAGGTTAAATGCTGGAATAATGCAAT 2050
Qy 2041 GTATTTCACTGCTCTTAAATCTTTGCTCTTGAATAACCAAGATCACTTGTCTAAATCT 2100
Db 2051 GTATTTCACTGCTCTTAAATCTTTGCTCTTGAATAACCAAGATCACTTGTCTAAATCT 2110
Qy 2101 ACTGTTTCTAGCTTGGACTATAGCAGAGATGAGTGGCTAGAGTCACTTTCCTGCTTAAAT 2160
Db 2111 ACTGTTTCTAGCTTGGACTATAGCAGAGATGAGTGGCTAGAGTCACTTTCCTGCTTAAAT 2170
Qy 2161 TTAAGTCACTGATGATGCTGAGACTCTATGCAACAAACATTTAAGGATTTATG 2220
Db 2171 TTAAGTCACTGATGATGCTGAGACTCTATGCAACAAACATTTAAGGATTTATG 2230
Qy 2221 AGAGCTAATGTTGAATCTTTAAATTTGGGAAATTTGAGTTGTTAGTCAACCCAGTCAAT 2280
Db 2231 AGAGCTAATGTTGAATCTTTAAATTTGGGAAATTTGAGTTGTTAGTCAACCCAGTCAAT 2290
Qy 2281 GATAAAAACAAACGATTTCTCTGAAGCTCTTGATATCTCGATGGAAGCATGTGAAGAC 2340
Db 2291 GATAAAAACAAACGATTTCTCTGAAGCTCTTGATATCTCGATGGAAGCATGTGAAGAC 2350
Qy 2341 AAGGCCAATCTTCATGCTCTCATTCAGATGAACCAAGGCTTATCCATCTCGATGGAAGCAT 2400
Db 2351 AAGGCCAATCTTCATGCTCTCATTCAGATGAACCAAGGCTTATCCATCTCGATGGAAGCAT 2410
Qy 2401 GCTTGTCTCTGCTGAGATTTCTCTCAATTCCTCAAGGATTTCTCTATCTGATGAA 2460
Db 2411 GCTTGTCTCTGCTGAGATTTCTCTCAATTCCTCAAGGATTTCTCTATCTGATGAA 2470
Qy 2461 AGAGGTTATGTAGCAAAACAAATTTGAAAGTGGCACAGGGAATCAAACTCCAAATATGTT 2520
Db 2471 AGAGGTTATGTAGCAAAACAAATTTGAAAGTGGCACAGGGAATCAAACTCCAAATATGTT 2530
Qy 2521 GACTTGTAGGAAACAACTCAATGAAGCACTTACTTACCGGAAGCCCTGTTGATGGT 2580
Db 2531 GACTTGTAGGAAACAACTCAATGAAGCACTTACTTACCGGAAGCCCTGTTGATGGT 2590
Qy 2581 GATAACTATGTTCTGCGAGTAAACCAAGATTAAGCGTCTCAGCTTACTGCTCTATA 2640
Db 2591 GATAACTATGTTCTGCGAGTAAACCAAGATTAAGCGTCTCAGCTTACTGCTCTATA 2650
Qy 2641 CACTTTTATGGAACACTAGTACACATATAACAGGCTGCCATTTGTTGGAAGTACAGAA 2700
Db 2651 CACTTTTATGGAACACTAGTACACATATAACAGGCTGCCATTTGTTGGAAGTACAGAA 2710
Qy 2701 ATTAFTACAGAACTCTGTGTAATGTTTCGTACACAGATTTGGATAAGTGGGAAGAAAT 2760
Db 2711 ATTAFTACAGAACTCTGTGTAATGTTTCGTACACAGATTTGGATAAGTGGGAAGAAAT 2770
Qy 2761 AAAAACTGAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAATTTGGGCTC 2820
Db 2771 AAAAACTGAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAATTTGGGCTC 2830

Qy 2821 AATTGCTACAGGAAGAAACGTTATCCAGATATATAAACTTGCAAAAACAGTGTGAA 2880
Db 2831 AATTGCTACAGGAAGAAACGTTATCCAGATATATAAACTTGCAAAAACAGTGTGAA 2890
Qy 2881 GTTCTTCCATCAGAGGACCTGTGTATATGTAATTTGGGCTCATAGCTAAACCAACAA 2940
Db 2891 GTTCTTCCATCAGAGGACCTGTGTATATGTAATTTGGGCTCATAGCTAAACCAACAA 2950
Qy 2941 GGCTGTGATATCTTAAATGTCAAACTGGGATGCTGTGAGGCATAGTCGCAAAACATCTG 3000
Db 2951 GGCTGTGATATCTTAAATGTCAAACTGGGATGCTGTGAGGCATAGTCGCAAAACATCTG 3010
Qy 3001 TGGCCAGTGGTTCAGAGATGATGTGGAACAACTCTGTAAATGAACCTTCTATATCCCAAGC 3060
Db 3011 TGGCCAGTGGTTCAGAGATGATGTGGAACAACTCTGTAAATGAACCTTCTATATCCCAAGC 3070
Qy 3061 ACTCTAAGTTTGAACCTGGAGTCAACAGCTCTAGACATAATAGTGAAGATGAATCTGTG 3120
Db 3071 ACTCTAAGTTTGAACCTGGAGTCAACAGCTCTAGACATAATAGTGAAGATGAATCTGTG 3130
Qy 3121 CCATCGAGTATGTTCAATTCGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3180
Db 3131 CCATCGAGTATGTTCAATTCGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACATTT 3190
Qy 3181 TTCTTGTATCAATGAAGATACAGAGCAACATTTTATGACCGATCTGGACCCCAATAAG 3240
Db 3191 TTCTTGTATCAATGAAGATACAGAGCAACATTTTATGACCGATCTGGACCCCAATAAG 3250
Qy 3241 GATAAAAATTCATCCCTTTCTTCTGTTCTAGTAAACCTTGTGGAAGATTCGTATCTTAAT 3300
Db 3251 GATAAAAATTCATCCCTTTCTTCTGTTCTAGTAAACCTTGTGGAAGATTCGTATCTTAAT 3310
Qy 3301 TCGCTTACTTTCCTTAAACAAACATCGTAGTAGCAGTATCCAAAGGAGGGAATTA 3360
Db 3311 TCGCTTACTTTCCTTAAACAAACATCGTAGTAGCAGTATCCAAAGGAGGGAATTA 3370
Qy 3361 TCATCTGAAAGTAAAGCAACAGGCGAATTCAGAACACTTTACGAGCCCAAGTGTGAT 3420
Db 3371 TCATCTGAAAGTAAAGCAACAGGCGAATTCAGAACACTTTACGAGCCCAAGTGTGAT 3430
Qy 3421 TTTAATCATAGTGTATGATTTTACACCCATATCCATCTGTACAGAAACATTTCAATAGAG 3480
Db 3431 TTTAATCATAGTGTATGATTTTACACCCATATCCATCTGTACAGAAACATTTCAATAGAG 3490
Qy 3481 ACTTCATTTATGGGGAATAAGCACATTTGAAGACACTTGTGTAGTACACCAAGCATTTGGAGA 3540
Db 3491 ACTTCATTTATGGGGAATAAGCACATTTGAAGACACTTGTGTAGTACACCAAGCATTTGGAGA 3550
Qy 3541 AATGACTTTAAATTCACCAAGAAATTTTGGTACAGAGAAATCAAGAGAAATACAGCCGA 3600
Db 3551 AATGACTTTAAATTCACCAAGAAATTTTGGTACAGAGAAATCAAGAGAAATACAGCCGA 3610
Qy 3601 GAGAGTTAGTGTAGTAAAGTTTCAACGAGCTCAATATGAAGATACGTAGCCAAAGTTTC 3660
Db 3611 GAGAGTTAGTGTAGTAAAGTTTCAACGAGCTCAATATGAAGATACGTAGCCAAAGTTTC 3670
Qy 3661 AATACAGACACTCAACAGTGGCATTAAGTTCAATGAGCTCAAGTCTTTCAGGAGACA 3720
Db 3671 AATACAGACACTCAACAGTGGCATTAAGTTCAATGAGCTCAAGTCTTTCAGGAGACA 3730
Qy 3721 GTAGTGTAGATGCTTACAACTATGGAACAGACTGTGGAAGCATGAGTACTGTGTAAGT 3780
Db 3731 GTAGTGTAGATGCTTACAACTATGGAACAGACTGTGGAAGCATGAGTACTGTGTAAGT 3790
Qy 3781 ACTAAAACTATTAAGCAAGCCACTATTTGACGCCACAGTCTAACCTCTCTCTCTCC 3840
Db 3791 ACTAAAACTATTAAGCAAGCCACTATTTGACGCCACAGTCTAACCTCTCTCTCTCC 3850
Qy 3841 AATCAAAATTCGGTGTCCCTGGTCCCTCCAGGTTCTTCTCATAGCTTCTCTAGAGAGCA 3900
Db 3851 AATCAAAATTCGGTGTCCCTGGTCCCTCCAGGTTCTTCTCATAGCTTCTCTAGAGAGCA 3910

3901 CAGTCCCTTAAAGCACCCCTCTATTGCTACAATTAATAAGCTCTAGCAGATTGTAACCTTTAGT 3960
3911 CAGTCCCTTAAAGCACCCCTCTATTGCTACAATTAATAAGCTCTAGCAGATTGTAACCTTTAGT 3970
3961 TACACAAGTCTAGAGATGCTTTGGCTATGCTACACGAAAGACTACAGCAACAAGA 4020
3971 TACACAAGTCTAGAGATGCTTTGGCTATGCTACACGAAAGACTACAGCAACAAGA 4030
4021 ATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGCTATT 4080
4031 ATGCATCCATCTTATCTCACTCTGAAGCTTTGGCATCTCCAGCAAAAGATGCTATT 4090
4081 ACTGATACCATACCATGAAGCCAAAGTTTGGATCAGATTACACCAAGCAGGTTTC 4140
4091 ACTGATACCATACCATGAAGCCAAAGTTTGGATCAGATTACACCAAGCAGGTTTC 4150
4141 ATGAAAGCCTTAAGTTATGATCATCATTAAGATAAAGAGATTATTAGTCTCTATTATCAA 4200
4151 ATGAAAGCCTTAAGTTATGATCATCATTAAGATAAAGAGATTATTAGTCTCTATTATCAA 4210
4201 AATACCTGCAACGATCTTCTCAGTCCGCTCCATGCTCCAGTCCCAATATGAGGGT 4260
4211 AATACCTGCAACGATCTTCTCAGTCCGCTCCATGCTCCAGTCCCAATATGAGGGT 4270
4261 TCAGATGATTACATTTGCTCTCTCCCGGTGGATATAAATGATATATTCAGAGTAAAG 4320
4271 TCAGATGATTACATTTGCTCTCTCCCGGTGGATATAAATGATATATTCAGAGTAAAG 4330
4321 GATATTCCTTATTTTACAGACAAAACATACACCATGATGATCGAGTGCAGAGCA 4380
4331 GATATTCCTTATTTTACAGACAAAACATACACCATGATGATCGAGTGCAGAGCA 4390
4381 TTTGCCATGATGAGGAGTCTTCCATCTGGAACTGGAGGCTTTGTAATAAATCTTTT 4440
4391 TTTGCCATGATGAGG----- 4407
4441 CACTTGCTACGACGAGATGAGTCTTACGGAATAATGAATTCATTCATTCAGATGCC 4500
4408 ----- 4407
4501 TCTCTGTTTTAGAAAGTACAGAGACACTGGACTACAGGAACATACAGATGATATGTC 4560
4408 -----AGAAAGTACAGAGACACTGGACTACAGGAACATACAGATGATATGTC 4456
4561 CTTTATGCTCTGTTAAATCTGGTTTCCAGCCAGCAACCACTGAGTGCATA 4620
4457 CTTTATGCTCTGTTAAATCTGGTTTCCAGCCAGCAACCACTGAGTGCATA 4516
4621 TGATGCTATTCAGATCTTCAAGATATTCATATTCATGTTGGTGTGAGCAGACTATCCAT 4680
4517 TGTAGTCAATTCAGATCTTCAAGATATTCATATTCATGTTGGTGTGAGCAGACTATCCAT 4576
4681 AATCCTTTAGAGTGGTTCCCTTAAGTTTTCGGGATTTCTGATGATGAGTGGGGT 4740
4577 AATCCTTTAGAGTGGTTCCCTTAAGTTTTCGGGATTTCTGATGATGAGTGGGGT 4636
4741 TCTCAAGAGCTCAGCTAGCAGCAACCAAGCAGAAATGTTACTAGTGTGTTAAACA 4800
4637 TCTCAAGAGCTCAGCTAGCAGCAACCAAGCAGAAATGTTACTAGTGTGTTAAACA 4696
4801 ATTCAGATGATACACCAATGTCGCTATCTCTCGCAAGAGTCTTAAGATTAGTC 4860
4697 ATTCAGATGATACACCAATGTCGCTATCTCTCGCAAGAGTCTTAAGATTAGTC 4756
4861 ATTAATTTAGTGTGATTTCAATTAATGTCATGAGCTGGCTTTTAAACAATTAG 4920
4757 ATTAATTTAGTGTGATTTCAATTAATGTCATGAGCTGGCTTTTAAACAATTAG 4816
4921 GAGAAGTATCCTCAACATTTGATGACATATGCTCTTACTCTGAGGTTTCCATTTGCTG 4980
4817 GAGAAGTATCCTCAACATTTGATGACATATGCTCTTACTCTGAGGTTTCCATTTGCTG 4876
4981 TCACACTGCAATTCAGACTTCCGCTGTCGGAGGTTTCATACAGAATTAATTTCAAGATGA 5040

Db 4877 TCACACTGCAATTCAGACTTCCGCTGCGAGGTTTCATACAGAATTAATTTCAAGATGA 4936
Qy 5041 CAGTTTCTACAAATGCAATGAAGAGCAGAGCTGTGTGGCAACACCAAGCAACCT 5100
Db 4937 CAGTTTCTACAAATGCAATGAAGAGCAGAGCTGTGTGGCAACACCAAGCAACCT 4996
Qy 5101 ATAGTTGATACATCTGCTGAATCCTGA 5127
Db 4997 ATAGTTGATACATCTGCTGAATCCTGA 5023

RESULT 2
HSK805608 4696 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA; cDNA DKFZp761C062 (from clone DKFZp761C062).
DEFINITION AL834497
ACCESSION AL834497.1 GI:21740290
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4696)
AUTHORS Ottenwaelder B., Obermaier B., Deutschenbaue S., Schaipe A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M. and Wiemann S.
CONSRM The German cDNA Consortium
TITLE Direct Submision
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaeder Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp761C062) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761C062
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
source

1. 4696
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761C062"
/db_xref="taxon:9606"
/clone="DKFZp761C062"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="hypothetical protein, N-terminus truncated, differentially spliced"
1. 4696
/gene="DKFZp761C062"
c3. 2408
/gene="DKFZp761C062"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD39155.1"
/db_xref="GI:21740291"
/db_xref="UniProt/TREMBL:Q8N3A0"
/translation="NIGSNWGLNLLQBEVVPDILKAKQCEVLSIRGTCVYVLGLI
AKTKQCDILKCHNWDVHRKHLWPVVDVQVLCNELSIPSTLSUNSESTSRH
NSESVPNSFLIEDDRFGSSSTFFLDINEDTEPTFYDRSGPKDKNSPPFFASS
KLVKNIILSLTLPNKKHSSSDPKGKLSSEKTSNRIRITLTFEVDNFHSDFTTP
ISTVQTLQLETSPMCKHLEDGTPSISGENDLFTKFTGNTENHRENTSRRLVVES
STSHWKIKSQSFNDITGSGISSMSSPSRRTVMDTDCGSMSTVYSTKTIK
TSHTLTPQNLHLSKSNVSLVPGSSHTLPRRAQSLKAPSIATIKSLADCNFSYTS
SRDAFGYATLRLQQRMPESLSHSEALSPAKVLDFTDTITTKANSFSRSLTPSRID
FKKHVGGIRSLRPTITNLPFKMALSALDKEDLSPINQITLQRSSSVSRMSVSS
ATVGGSDDYIGLALPVDINDIFQVQKIPYFQTKNIPPHDDRGARAFADAGGLPSGTG

GLVKNFLLRQMSLBEIMNSIHSASLFLSESTBDTGLQEHDDNCLYCVIEILGF
QPSNLSAICSHSDFDPIPYDWCSEQTHNPLEVVPSPSGISGCSGVSEGSAST
KSTELLGKVPIDPTMPCRIILRKVELRLVINLSSSVSKTCHETGLLTIKPKYPTF
DDICLYSVSHLLSHCTPRLPCRRPIQLBFDQVLOPQWHEBEAEVLAFLPPKQPIVDS
AES"

ORIGIN

Query Match 44.0%; Score 2254; DB 3; Length 4696;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2336; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

Qy	2792	GAATATCGGCTCATCAATTTGGGCTCTCAATTTGCTACAGGAAGAAACGTGATCCAG	2851
Db	1	GAATATCGGCTCATCAATTTGGGCTCTCAATTTGCTACAGGAAGAAACGTGATCCAG	60
Qy	2852	ATATACTAAACTTGCACAAACAGTGTGAAGTCTTTCCATCAGAGGACCTGTGTATATG	2911
Db	61	ATATACTAAACTTGCACAAACAGTGTGAAGTCTTTCCATCAGAGGACCTGTGTATATG	120
Qy	2912	TACTTGGGCTCATAGCTAAACCAAAACAAAGGCTGTGATATCTAAATGTCACAACTGGG	2971
Db	121	TACTTGGGCTCATAGCTAAACCAAAACAAAGGCTGTGATATCTAAATGTCACAACTGGG	180
Qy	2972	ATGCTGTGAGCAGTAGTCGCAACATCTGTGCGCAGTGTTCAGATCATGTGGAACAAC	3031
Db	181	ATGCTGTGAGGAGTAGTCGCAACATCTGTGCGCAGTGTTCAGATCATGTGGAACAAC	240
Qy	3032	TCTGTAATGAACCTTTCATCTATCCCAAGCAGCTTAAGTTTGAACCTCGGAGTCAACAGCT	3091
Db	241	TCTGTAATGAACCTTTCATCTATCCCAAGCAGCTTAAGTTTGAACCTCGGAGTCAACAGCT	300
Qy	3092	CTAGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTCATATTTGAGGATGACC	3151
Db	301	CTAGACATAATAGTGAAGTGAATCTGTGCCATCGAGTATGTTCATATTTGAGGATGACC	360
Qy	3152	GGTTTGGCAGAGCTCTACTAGTACATTTTCCCTTGATATCAATGAAGATACAGGCCAA	3211
Db	361	GGTTTGGCAGAGCTCTACTAGTACATTTTCCCTTGATATCAATGAAGATACAGGCCAA	420
Qy	3212	CATTTTATGACCGATCTGGACCCATAAGGATAAAATTTCAATCCCTTCTTTGCTCTTA	3271
Db	421	CATTTTATGACCGATCTGGACCCATAAGGATAAAATTTCAATCCCTTCTTTGCTCTTA	480
Qy	3272	GTAACCTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAAACAAAACATCGTA	3331
Db	481	GTAACCTTGTGAAGATCGTATCTTAAATTCGCTTACTTTGCTTAAACAAAACATCGTA	540
Qy	3332	GTACAGTGTATCCAAAGAGGGGAAATTTATCATCTGAAAGTAAAGCAAGCAAGCGGAA	3391
Db	541	GTACAGTGTATCCAAAGAGGGGAAATTTATCATCTGAAAGTAAAGCAAGCAAGCGGAA	600
Qy	3392	TCAGAACACTTTACGGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACCCCATAT	3451
Db	601	TCAGAACACTTTACGGAGCCAGTGTGATTTTAAATCATAGTATGATTTTACCCCATAT	660
Qy	3452	CCACTGTACGAAACATTAACATTTAGAGACTTCAATTTATGGGAAATGAAGCAATGAAG	3511
Db	661	CCACTGTACGAAACATTAACATTTAGAGACTTCAATTTATGGGAAATGAAGCAATGAAG	720
Qy	3512	ACACTGTAGTACACCAAGCAATTCGAGAAATGATTTAAATTCACCAAGAAATTTGGTA	3571
Db	721	ACACTGTAGTACACCAAGCAATTCGAGAAATGATTTAAATTCACCAAGAAATTTGGTA	780
Qy	3572	CAGAGATTCACAGAGAAATAACAGCCGAGAGAGGTAGTAGTAGAAGTTTCAACAGACT	3631
Db	781	CAGAGATTCACAGAGAAATAACAGCCGAGAGAGGTAGTAGTAGAAGTTTCAACAGACT	840
Qy	3632	CACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACCAACAGTGGCATAAGTT	3691
Db	841	CACATATGAAGATACGTAGCCAAAGTTTCAATACAGACACTACCAACAGTGGCATAAGTT	900
Qy	3692	CAATGAGCTCAAGTCTTTCAGAGAGACAGTGTGTAGATGCTTCAACATATGACACAG	3751
Db		CAATGAGCTCAAGTCTTTCAGAGAGACAGTGTGTAGATGCTTCAACATATGACACAG	

Db	901	CAATGAGCTCAAGTCTTTCACGAGAGACAGTAGTGTAGATGCTTACAACTATGGACACAG	960
Qy	3752	ACTGTGGAAGCATAGTACTGTGTAACTACTTAAACATTAATAAGACACAGCCACTATTGCA	3811
Db	961	ACTGTGGAAGCATAGTACTGTGTAACTACTTAAACATTAATAAGACACAGCCACTATTGCA	1020
Qy	3812	CGCCACAGTCTAACCCATCTGTCTCTCCAAATCAAATTCGGTGTCCCTTGGTCCCTCCAG	3871
Db	1021	CGCCACAGTCTAACCCATCTGTCTCTCCAAATCAAATTCGGTGTCCCTTGGTCCCTCCAG	1080
Qy	3872	GTTCTTCTCATACCGTCTCTAGAGAGACAGTCCCTTAAAGCACCCCTCTATTGTCTACAA	3931
Db	1081	GTTCTTCTCATACCGTCTCTAGAGAGACAGTCCCTTAAAGCACCCCTCTATTGTCTACAA	1140
Qy	3932	TTAAAGTCTAGCAGATTGTAACCTTGTACACAAGTTCTAGAGATGCTTTTGGCTATG	3991
Db	1141	TTAAAGTCTAGCAGATTGTAACCTTGTGTACACAAGTTCTAGAGATGCTTTTGGCTATG	1200
Qy	3992	CTACACTGAAAGACTACAGCAACAAAGAAATGCAATCCATCTTATCTCTGAAAGCTT	4051
Db	1201	CTACACTGAAAGACTACAGCAACAAAGAAATGCAATCCATCTTATCTCACTCTGAAGCTT	1260
Qy	4052	TGGCATCTCCAGCAAAAGATGTCTATTACTGATACCATCAACATGAAGGCCAACAGTT	4111
Db	1261	TGGCATCTCCAGCAAAAGATGTCTATTACTGATACCATCAACATGAAGGCCAACAGTT	1320
Qy	4112	TTGAGTCCAGATTAAACACCAAG-----	4133
Db	1321	TTGAGTCCAGATTAAACACCAAGAGGATCGATTTTAAAGAGCATGTCCGGGGAATCA	1380
Qy	4134	-----CAGGTTTCATGAAAGCCCTTAAAGTTATG	4159
Db	1381	GGAGCTTAAGACCTTACAATAACAAACCACTTTTTCAGGTTTCATGAAAGCCCTTAAAGTTATG	1440
Qy	4160	CATCATTAGATAAAGAGATTTTATTTAGTCTCTTATTAATCAAAATACCCCTGCAACGATCTT	4219
Db	1441	CATCATTAGATAAAGAGATTTTATTTAGTCTCTTATTAATCAAAATACCCCTGCAACGATCTT	1500
Qy	4220	CCTCAGTCCGCTCCATCGGTGTCAGTCCACATATGGGGTTTCAGATGATTTTGGTCTC	4279
Db	1501	CCTCAGTCCGCTCCATCGGTGTCAGTCCACATATGGGGTTTCAGATGATTTTGGTCTC	1560
Qy	4280	TTGCTCTCCGGTGGATATAAATGATATATTTCCAGGTAAAGGATATTCCTCTATTTCAGA	4339
Db	1561	TTGCTCTCCGGTGGATATAAATGATATATTTCCAGGTAAAGGATATTCCTCTATTTCAGA	1620
Qy	4340	CAAAAAACATACACCAACATGATGATGAGGTGCAAGAGCATTTGGCCCATGATGAGGAG	4399
Db	1621	CAAAAAACATACACCAACATGATGATGAGGTGCAAGAGCATTTGGCCCATGATGAGGAG	1680
Qy	4400	GTCTTTCCATCTGGAACCTGGAGGCTCTGTAAATAATCTTTTCACTTGTCTACGACAGCAGA	4459
Db	1681	GTCTTTCCATCTGGAACCTGGAGGCTCTGTAAATAATCTTTTCACTTGTCTACGACAGCAGA	1740
Qy	4460	TGAGTCTTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTCTGTTTAAAGAGTA	4519
Db	1741	TGAGTCTTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTCTGTTTAAAGAGTA	1800
Qy	4520	CAGAGACACTGGAATACAGAACATACAGATGATACTGCTTTTATTTGTCTGTATTTG	4579
Db	1801	CAGAGACACTGGAATACAGAACATACAGATGATACTGCTTTTATTTGTCTGTATTTG	1860
Qy	4580	AAATTTCTGGGTTTCCAGCCCAACCAACTGAGTGAATATGTAGTCAATTCAGACTTTC	4639
Db	1861	AAATTTCTGGGTTTCCAGCCCAACCAACTGAGTGAATATGTAGTCAATTCAGACTTTC	1920
Qy	4640	AAGATATTCATATTCATTTGTTGTTGAGCAGACATATCCATAATCTTTTAAAGTGTTC	4699
Db	1921	AAGATATTCATATTCATTTGTTGTTGAGCAGACATATCCATAATCTTTTAAAGTGTTC	1980
Qy	4700	CCTCTAAGTTTTCGGGATTTCTGGATGCAAGTGTGGGTGTCTCAAGAGGCTCAGCTA	4759
Db	1981	CCTCTAAGTTTTCGGGATTTCTGGATGCAAGTGTGGGTGTCTCAAGAGGCTCAGCTA	2040

QY	4760	GCAGCACCAAAAGCAGCAAGATTGTTACTAGGTGTTAAACAATTCAGATGATACACCA	4819
Db	2041	GCAGCACCAAAAGCAGCAAGATTGTTACTAGGTGTTAAACAATTCAGATGATACACCA	2100
QY	4820	TGTGCGGTACTCTCCCAAGAGTCTTAAGATTAGTCAATTAATTTGAGTGTTCAG	4879
Db	2101	TGTGCGGTACTCTCCCAAGAGTCTTAAGATTAGTCAATTAATTTGAGTGTTCAG	2160
QY	4880	TTTCAACTAAATGTCATGAGCTGGGCTTTTAAACAATTAAGGAGAGTATCTCTCAAAACAT	4939
Db	2161	TTTCAACTAAATGTCATGAGCTGGGCTTTTAAACAATTAAGGAGAGTATCTCTCAAAACAT	2220
QY	4940	TTGATGACATATGCTTACTCTGAGGTTTCCCATTTGCTGTGCACACTGCACATTCAGAC	4999
Db	2221	TTGATGACATATGCTTACTCTGAGGTTTCCCATTTGCTGTGCACACTGCACATTCAGAC	2280
QY	5000	TTCCGTTGCGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTTACAAATCGATG	5059
Db	2281	TTCCGTTGCGAGGTTTCATACAGAAATTTTCAAGATGTACAGTTTCTTACAAATCGATG	2340
QY	5060	AAGAAGCAGAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTG	5119
Db	2341	AAGAAGCAGAGGCTGTGTTGGCAACACCAACCAAGCAACCTATAGTTGATACATCTGCTG	2400
QY	5120	AATCCTGA 5127	
Db	2401	AATCCTGA 2408	
RESULT 3			
AK036149			
LOCUS			
DEFINITION	AK036149	2497 bp mRNA linear HTC 03-APR-2004	
		Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length	
		enriched library, clone:9630041G24 product:hypothetical ARM repeat	
		structure containing protein, full insert sequence.	
ACCESSION	AK036149		
VERSION	AK036149.1	GI:26331181	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	1	Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE			
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
		Itou, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
		Normalization and subtraction of cap-trapper-selected cDNAs to	
		prepare full-length cDNA libraries for rapid discovery of new genes	
		Genome Res. 10 (10), 1617-1630 (2000)	
		20499374	
		11042159	
REFERENCE			
AUTHORS	3	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,	
		Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,	
		Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,	
		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,	
		Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,	
		Yoneda, Y., Ishikawa, I., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,	
		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
		RIKEN integrated sequence analysis (RISA) system--384-format	
		sequencing pipeline with 384 multicapillary sequencer	
		Genome Res. 10 (11), 1757-1771 (2000)	
		20530913	
		11076861	
REFERENCE			
AUTHORS	4	The RIKEN Genome Exploration Research Group Phase II Team and the	
		FANTOM Consortium.	

TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6
AUTHORS	(bases 1 to 2497)
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/
FEATURES	Location/Qualifiers
source	1. 2497
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM DB:9630041G24"
	/db_xref="taxon:10090"
	/clone="9630041G24"
	/tissue_type="cerebellum"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="16 days neonate"
	24. .>2495
CDS	/note="unnamed protein product; hypothetical ARM repeat structure containing protein (SCOP 48371, evidence: SCOP putative"
	/codon_start=1
	/protein_id="BAC29321.1"
	/db_xref="GI:26331182"
	/translation="MAAIGRSLKRLRIRGRNDSSENVPLDITREPDSNLRILQNGRALRYLQDSILQVKLVKVDYLARCDIQNSVEVERQALVRMTVNASL VAKLOGVNNRKLHNLNPKLIDGHSSEELQFNEDYDIIICLRALNKEKVRVA FSSVANSIANGNDLQERDMRACIAICELALQNPVVVALRGANTILNKVIDC QLRNEALITLILHLPKTRQVYRADERILAPYTDHYRHSPTAEQLKDE FRPLPVPVDFEIALSVDPGRQSWLSDFVAEAKTILPHARSRPDLMDYL ALITSAIFRNGLEGLVEITNSDDHSIVRATILLGELHLMANTILPHSGSHLHCIP TLMNMAASPDIPKEKLRASALCLNRHFMKKRGPYSLHDHIIIOKATNHRK DOYLRVQKDI FVLKDTTEALLNDRSQVLQKENDLMDWNLGITLKPWNLRNYK DEQLHRFVRLLYFYKPSKLYASLDLDAKSKOLTVGCOFTFELSGEEDQGVLE DLVQDIIVQMLNASHVYKPSRQNGLLTSLQHVFLGTSLCHPGVKMLKCSVF QCCLNLCSLKNQDHLIKLVSSLDYSRDLGRLVILSKILTATDACLATYKHLVL RANVEFFNNWNGIELLYQLHDKNKTISSEALDILDEACEDKANLHIALIQKRPALSHLG DKGLLLLLRFLSPKGFSLNERYV"
ORIGIN	
Query Match	41.4%; Score 2123.6; DB 3; Length 2497;

Best Local Similarity 91.1%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

Qy	1	ATGCGGCGATCGGCGGCGCGCTCTCTGAAGAACCTCCGAGTACGAGGCGGGAATCAC	60	Db	1044	GTGTGACGGATGAGTTTCATAGAACCCCTGCTCAGTGTAGATCCAGGAAGATTCCAGGAC	1103
Db	24	ATGCGGCGATCGGCGGCGCGCTCTCTGAAGAACCTCCGATACGAGGCGGGAATCAC	83	Qy	1081	AGTTGGAGGCTTTCAGATGGCTTTGTGGCAGCTGAGGCAAAAACCTATTCTTCCCTCATCGT	1140
Qy	61	AGCGCGAGGAGAACGTCGCGCTGGATCTGACCCGAGAACCTTCTGATAAATTAAGAGAG	120	Db	1104	AGTTGGAGGCTTTCAGATGGCTTTGTGGCAGCTGAGGCAAAAACCTATTCTTCCCTCATCGT	1163
Db	84	AGCGCGAGGAGAACGTCGCGCTGGATCTGACCCGAGAACCTTCTGATAAATTAAGAGAG	143	Qy	1141	GCCAGATCCAGGCGCAGACCTCATGGATAATTAATTGGCAGCTGATCTCTGCAATTTAT	1200
Qy	121	ATTTCTCCAAATGTGGCAGATTCGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG	180	Db	1164	GCTAGATCCAGGCGCAGACCTCATGGACAAATTAATTGGCTCTGATACCTCTGCAATTTAT	1223
Db	144	ATTTCTCCAAATGTGGCAGATTCGAGGAGTATCAAAATATGAGAAAGCTAGGCCATCTG	203	Qy	1201	CGTAATGGACCTTTTAGAGGCTCTAGTGAAGTATTAACAAACAGTGAATCATATCTCA	1260
Qy	181	AATAACTTTACTAAGCTTTCTTGATATTTGGCCACAGTGAAGAAAAAACTGGCTTTCCAC	240	Db	1224	CGTAATGGACCTTTTAGAGGCTTGGTCGAGGTGATAACGAAACAGTGAACACATCTCA	1283
Db	204	AATAACTTTACTAAGCTTTCTTGATATTTGGCCACAGTGAAGAAAAAACTGGCTTTCCAC	263	Qy	1261	GTTAGAGCTACCATCTCTTTTAGGAGAGCTTTTACATATAGCAAAACACAAATTTCTTCTCAT	1320
Qy	241	TATGAGGATATCATAAATTTGTTGCGTTAGCTTTTAAATGAAGCAAAAAGAAAGTGCAG	300	Db	1284	GTGAGGCGGACCATCTCTTTTAGGAGAGCTCTGATATTTGGCAAAACACCATCTGCTCAC	1343
Db	264	TATGAGGATATCATAAATTTGTTGCGTTAGCTTTTAAATGAAGCAAAAAGAAAGTGCAG	323	Qy	1321	TCACATAGCCATCATTTACACTGCTTGCACACCTCAATGAATATGGCTGCAATCTTTGAT	1380
Qy	301	GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTTCTCCAGAGGTG	360	Db	1344	TCGATAGCCACCATCTACACTGCTTGCACACCTCATGAACATGGCTGCGTCTTCGAT	1403
Db	324	GCAGCAGGCTACGAGCGCTTCGATATCTCATCCAGACTCCAGTATTTCTCCAGAGGTG	383	Qy	1381	ATCCCAAGGAAAGAGAGCTGCGAGCCAGTGCAGCCCTGAACTGTTTGAACCGCTTCCAT	1463
Qy	361	CTAAATTTGAAGTGGACATTTAATAGCTAGTGTGATTTGATACATACACAGCAACGAG	420	Db	1404	ATTTCCCAAGGAAAGAGAGCTGCGAGCCAGTGCAGCCCTGAACTGTTTGAACCGCTTCCAT	1463
Db	384	CTAAATTTGAAGTGGACATTTAATAGCTAGTGTGATTTGATACATACACAGCAACGAG	443	Qy	1441	GAATGAAGAAACAGGAGCCTTAAGCCTTATAGTCTTCAATTTAGACCAATTTATTCAGAAA	1500
Qy	421	GTAGAGGAGCACAAAGCATTTCGATTTAGTCAGAAAGATGATTACTGTGAATGCTTCCCTTG	480	Db	1464	GAGATGAAGAAACAGGAGCCTTAAGCCTTATAGTCTTCAATTTAGACCAATTTATTCAGAAA	1523
Db	444	GTGAGAGGACACAAGCCCTTCGTTAGTCAGAAAGATGATTACTGTGAATGCTTCCCTTG	503	Qy	1501	GCAATTCGCAACACACCAAGGAGTTCAGGATCTCCGAGTTTCAGAAAGATATATTTATC	1560
Qy	481	TTTCTAGTCTGACCACTCATTAATTCAGTTGGAATGGAATGATGAGACTTCAAGAAAGA	540	Db	1524	GCAATTCGCAACACACCAAGGAGTTCAGGATCTCCGAGTTTCAGAAAGATATATTTGTT	1583
Db	504	TTTCTAGTCTGACCACTCATTAATTCAGTTGGAATGGAATGATGAGACTTCAAGAAAGA	563	Qy	1561	CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAGTCTTCAACAT	1620
Qy	541	GACAGATGGTCCGAGCATGATTCGCAATTAATTCGTAAGTACAGTCTTCAAGATCCAGAG	600	Db	1584	CTTAAGGATACAGAGGAGCTCTTTTAAATTAACCTTAGAGATAGCCAGTCTTCAACAT	1643
Db	564	GACAGATGGTCCGAGCATGATTCGCAATTAATTCGTAAGTACAGTCTTCAAGATCCAGAG	623	Qy	1621	AAAGAGAACTTTGAATGGAACTTTATAGGAGCCTTCTTAAGTGGCCAAATGTA	1680
Qy	601	GTGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAATATGTAATGTTGCAATTA	660	Db	1644	AAAGAGAACTTTGAATGGAACTTTATAGGAGCCTTCTTAAGTGGCCAAATGTA	1703
Db	624	GTGTGGCCCTTCGAGGAGGACTAAACACCATATTTGAAATATGTAATGTTGCAATTA	683	Qy	1681	AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT	1740
Qy	661	AGTCGAATAAATGAGGCTTAATTAATCAAAATTTTGCACTTCTTAATCATCCAAAGACT	720	Db	1704	AATCTAAGAACTATAAAGATGAACAGTTACACAGGTTTGTACGAAGACTACTTTATTTT	1763
Db	684	AGTCGAATAAATGAGGCTTAATTAATCAAAATTTTGCACTTCTTAATCATCCAAAGACT	743	Qy	1741	TACAGCCCGAGAGTAAATTAATATGCGCAACCTGATCTGGAATTTTGGCAAGGCGCAACAG	1800
Qy	721	CGGAGATATGTCGAGCTGATGTAAGTATAGAGAAATTTAGCACCTTACTGATTTT	780	Db	1764	TACAGCCCGAGAGTAAATTAATATGCGCAACCTGATCTGGAATTTTGGCAAGGCGCAACAG	1823
Db	744	CGGAGATATGTCGAGCTGATGTAAGTATAGAGAAATTTAGCACCTTACTGATTTT	803	Qy	1801	CTCACGGTTCTGAGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAGGAGTGGGCAA	1860
Qy	781	CACATACAGACATAGTCAGATACAGCTCAAGGACAGCTCAAGAAAGACAGAGAGCAGA	840	Db	1824	CTCACGGTTCTGAGTTGCCAGTTTACAGAAATTTCTTCTTGAATCTGAGGAGTGGGCAA	1883
Db	804	CACATACAGACATAGTCAGATACAGCTCAAGGACAGCTCAAGAAAGACAGAGAGCAGA	863	Qy	1861	GGCTACTTTAGAGATCTAGTAAAGGATATTTGTCAGTGGCTCAATGCTTCACTGGAGTG	1920
Qy	841	TTTCTAGCGATAAATGGGAATCATAGCAACATTTCCGATCATGGCAGGATTTATTAAT	900	Db	1884	GGATACTTTAGAGATCTAGTAAAGGATATTTGTCAGTGGCTCAATGCTTCACTGGAGTG	1943
Db	864	TTTCTAGCGATAAATGGGAATCATAGCAACATTTCCGATCATGGCAGGATTTATTAAT	923	Qy	1921	AAACCCGAAAGAGTCTTCAAAATAATGTTTATGACCAACCTTTAGTCAACACTACTTTT	1980
Qy	901	TTATGTAACCTGGAATTTCTGGATCCAGTCTTAATAGAGTACTTTCGATACCAAT	960	Db	1944	AAACCTGAGAGAGGCTTTCAGAAATTAATGGCTTACTGACCCCTTTAGTCAACACTACTTC	2003
Db	924	TTATGTAACCTGGAATTTCTGGATCCAGTCTTAATAGAGTACTTTCGATACCAAT	983	Qy	1981	TTATTTATTTGGAACACTTTCTTCCGACCTTCAGTGGAGTTAAATGCTGGAAAAATGCAGT	2040
Qy	961	ATGGAATAAGGCGAGGCTCTGCTCGAAGTGTCTATGACATCTTTGCTCTCCCTCTGCC	1020	Db	2004	TTATTTATTTGGAACACTTTCTTCCGACCTTCAGTGGAGTTAAATGCTGGAAAAATGCAGT	2063
Db	984	ATGGAATAAGGCGAGGCTCTGCTCGAAGTGTCTATGACATCTTTGCTCTCCCTCTGCC	1043	Qy	2041	GTATTTTCAAGTGTCTCTTAAATCTTTTGTCTTCTTGAATAACCAAGATCACTTGTGTAACCTT	2100
Qy	1021	GTGTGATGAGGAGTTCTAGAAAGCACTACTCAGTGTAGATCCAGGAGGTTCCAAAGAC	1080	Db	2064	GTCTTTTCAAGTGTCTCTTAAATCTTTTGTCTTCTTGAATAACCAAGATCACTTGTGTAACCTT	2123
Db				Qy	2101	ACTGTTTCTAGTGTGGACTATAGCAGAGTGGATTTGGCTAGAGTCACTCTTCTTCAAAAT	2160
				Db	2124	ACTGTTTCTAGTGTGGACTATAGCAGAGTGGATTTGGCTAGAGTCACTCTTCTTCAAAAT	2183

```

QY 2161 TTAACGTGAGCTACTGATGCTCCAGACTCTATGCAACAAACATTTAAGGTATTATG 2220
Db 2184 CTCACGGGCACACAGATCTTGACAGCTGTATGCAACAAACATTTAAGGTATTATG 2243
QY 2221 AGAGCTAAATGTTGAATCTTTTAATAATTTGGGAAATTTAGTCTGTAGTACCCAGCTACAT 2280
Db 2244 AGAGCTAAGTTGAATCTTTCAACAAATTTGGGAAATTTAGTCTGTAGTACCTACGCTACAT 2303
QY 2281 GATAAAACAAACAGATTTCTCTGAAGCTCTTGATATCTCTGATGAGCATGTGAAGAC 2340
Db 2304 GATAAAACAAACAAATTTCTCTGAAGCTCTTGATATCTCTGATGAGCATGTGAAGAC 2363
QY 2341 AAGGCCAATCTTCATGCTCTCATTCAGATGAACACGCTTATCCACCTTGGAGCACAAG 2400
Db 2364 AAGGCCAATCTTCATGCTCTCATTCAGATGAACACGCTTATCCACCTTGGAGCACAAG 2423
QY 2401 GGTTCCTCTCTGCTGAGATTTCTCTCCATTCCAAAGAGATTTTCCTATCTGAATGAA 2460
Db 2424 GGTTCCTCTCTGCTGAGATTTCTCTCCATTCCAAAGAGATTTTCCTATCTGAATGAA 2483
QY 2461 AGAGTTATGTAGC 2474
Db 2484 AGGGTTATGTGC 2497

RESULT 4
AK082586
LOCUS
DEFINITION
Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:C230066117 product:hypothetical ARM repeat
structure containing protein, full insert sequence.
ACCESSION
AK082586
VERSION
AK082586.1 GI:26100768
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
AUTHORS
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
AUTHORS
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the
AUTHORS
PANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
MEDLINE
11076861
PUBMED
11076861
REFERENCE
5 The PANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS

```

```

Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
6 (bases 1 to 3881)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saitoh, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, I., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
source
1..3881
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="PANTOM DB:C230066117"
/db_xref="taxon:10090"
/clone="C230066117"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
misc_feature
24..2581
/note="hypothetical ARM repeat structure containing
protein (SCOP|48371, evidence: SCOP)
putative"
3868..3873
/note="putative"
polyA_signal
3881
/note="putative"
polyA_site
3881
/note="putative"
ORIGIN
Query Match 41.3%; Score 2116.8; DB 3; Length 3881;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 2270; Conservative 0; Mismatches 222; Indels 10; Gaps 1;
QY 1 ATGGCGGCGATCGCGCGCGCGCTCTCTGAGAAACCTCCGAGTACGAGCGCGGATGAC 60
Db 24 ATGGCGGCGATCGCGCGCGCGCTCTCTGAGAAACCTCCGAGTACGAGCGCGGATGAC 83
QY 61 AGCGGCGAGGAGACGTCCTCGCTGGATCTGACCCGAGAACCTTCTGATTAACATTAAGAG 120
Db 84 AGCGGCGAGGAGACGTCCTCGCTGGATCTGACCCGAGAACCTTCTGATTAACATTAAGAG 143
QY 121 ATTCTCCAAATGTGGCCAGATTGAGGAGGATCAATATGAGAAAGCTAGGCCATCTG 180
Db 144 ATTCTCCAAATGTGGCCAAATTTGAGAGGATCAATATGAGAAAGCTAGGCCATCTG 203
QY 181 AATAACTTACTTAAGCTCTTTGTGATATTGGCCACAGTGAAGAAAAAATCTGGCTTTTAC 240
Db 204 AATAACTTACTTAAGCTCTTTGTGATATTGGCCATAGTGAAGAAAAAATCTGGCTTTTAC 263

```

Qy 241 TATGAGGATATCATAATTTGTTGCGGTAGCTTTATTAATGAAGCAAAAGAGTGCGA 300
Db |||||
Qy 264 TATGAAGACATCAATAATTTGTTGCGCTTAGCTTTATTAATGAGGCAAAAGAGTAAGA 323
Db |||||
Qy 301 GCAGCAGGCTACGAGCGCTTCGATATCTCATCAAGACTCAGTATTTCTCCAGAGGTG 360
Db |||||
Qy 324 GCAGCGGGTTGCGAGCGCTCGGTATCTCATCAAGATTCCAGTATTTCTCCAGAGGTG 383
Db |||||
Qy 361 CTAATAATGAAGTGGACTATTTAATAGCTAGGTGCTTTGATACATACACAGCAACGAG 420
Db |||||
Qy 384 CTAATACTGAAGTGGACTATTTAATAGCTAGGTGCTTTGATTTAGCAGAGCAACGAG 443
Db |||||
Qy 421 GTAGAGGACACAAGCACTTCGATTTAGTCAGAAGAGATGATTACTGTGAATGCTTCCTTG 480
Db |||||
Qy 444 GTGAGAGGACACAAGCCCTTCGCTTAGTCAGAAAGATGATTAATGTGAATGCTTCCTTG 503
Db |||||
Qy 481 TTTCTAGTTCTGTGACCACTCATTAATTCAGTTGGAATGATGACTTCAAGAAGA 540
Db |||||
Qy 504 TTTCTAGTTCTGTGCGCACTCATTAATTCGCTTGGAAATGATGGCTTCAGGAACGA 563
Db |||||
Qy 541 GACAGAATGTCGAGCATGATTGCAATTTATCTGTGAATAGCACTTCAGAAATCCAGAG 600
Db |||||
Qy 564 GACAGAATGTCGAGCTTGCTGCGCATCATCTGTGAATAGCACTTCAGAACCCAGAG 623
Db |||||
Qy 601 GTGTGCGCTTCGAGGAGGACTAAACACCATATTTGAAATGTTGATTGTCCAATTA 660
Db |||||
Qy 624 GTGTGCGCTTCGAGGAGGACTAAACACCATATTTGAAATGTTGATTGTCCAATTA 683
Db |||||
Qy 661 AGTCGAATTAATGAGGCTTAATTAACAATTTTGCACTTCTTAATCATATCAATCAAGACT 720
Db |||||
Qy 684 AGTCGAATTAATGAGGCTTCAATTAACAATTTTGCACTTCTTAATCATATCAATCAAGACT 743
Db |||||
Qy 721 CGGAGATGTCGAGCTGATGAGAATTTAGAGAGATTTTAGCACCTTACTGATTTT 780
Db |||||
Qy 744 CGGAGATGTCGAGCTGATGAGAATTTAGAGAGATTTTAGCACCTTACTGATTTT 793
Db |||||
Qy 781 CACTACAGACATAGTCAGATACAGCTGAAGGACAGCTCAAAAGAGACAGAGACGCA 840
Db |||||
Qy 794 CACTACAGACATAGTCAGATACAGCTGAAGGACAGCTCAAAAGAGACAGAGACGCA 853
Db |||||
Qy 841 TTTCTAGCGAGTAAATGCGAATCATAGCAACATTCGATCATCGGAGGAGTATTAAT 900
Db |||||
Qy 854 TTTCTAGCGAGTAAATGCGAATCATAGCAACATTCGATCATCGGAGGAGTATTAAT 913
Db |||||
Qy 901 TTATGTAACCTGGAATTTCTGGATCCAGTCTCTAATAGGAGTACTTTGCATACCAAT 960
Db |||||
Qy 914 TTATGTAACCTGGAATTTCTGGATACAGTCTCTAATGGAGTACTTTGCATACCGAAT 973
Db |||||
Qy 961 ATGGAATAAGGCGAGGTCTACTTGAAGTGTCTTTATGATATATTTCTCTCTACCT 1020
Db |||||
Qy 974 ATGGAATAAGGCGAGGTCTCTGAGTGTCTCTATGACATCTTTCTCTCTCTGCCC 1033
Db |||||
Qy 1021 GTTGTGATGAGGAGTCTATAGAGACACTACTAGTGTAGATCCAGGAGGTTCCAGAG 1080
Db |||||
Qy 1034 GTTGTGATGAGTCTATAGAGACCTGCTCAGTGTAGATCCAGGAAGATTCAGGAC 1093
Db |||||
Qy 1081 AGTTGAGGCTTTCAGATGGCTTTGTCGAGCTGAGGCAAAACTATTTCTCTCATCGT 1140
Db |||||
Qy 1094 AGTTGAGGCTTTCAGATGGCTTTGTCGAGCTGAGGCAAAACTATTTCTTCCCATCGT 1153
Db |||||
Qy 1141 GCCAGATCCAGGCGAGACCTCATGGAATTAATTTGGCACTGATCTCTGCAATTTAT 1200
Db |||||
Qy 1154 GCTAGATCCAGGCGAGACCTCATGGAATTAATTTGGCTCTGATCTCTGCAATTTAT 1213
Db |||||
Qy 1201 CGTAATGCACTTTTGAAGGCTTGTAGTGAAGTATTAACAAAGTGTATCATATCTCA 1260
Db |||||
Qy 1214 CGTAATGCACTTTTGAAGGCTTGTGTGAGTATTAACAAAGTGTATCATATCTCA 1273
Db |||||
Qy 1261 GTTAGACTACCATCTTTTAGGAGCTTTTACATATGGCAACAAATTTCTTCTCAT 1320
Db |||||
Qy 1274 GTGAGGCGACCATCTCTTTAGGAGCTCTCTGATATGGCAACAAATCTCTGCTCAC 1333
Db |||||

Qy 1321 TCACATAGCCATCATTTACACTGCTTGCCAAACCTTAATGAATATGCTGCTCTTTGAT 1380
Db |||||
Qy 1334 TCGCATAGCCACCATCTACACTGCTTGCCAAACCTCATGAACATGCTGCTCTTCAT 1393
Db |||||
Qy 1381 ATCCCAAGGAAAGAGACTGCGAGCGAGTGCAGCTTTGAACTGTTTAAACGCTTCAT 1440
Db |||||
Qy 1394 ATCCCAAGGAAAGAGGCTGCGAGCGAGTGCAGCTTTGAACTGTTTAAACGCTTCAT 1453
Db |||||
Qy 1441 GAAATGAAGAAACAGAGCCTTAAGCCTTTATAGTCTTCAATTTAGACCAATTTTACAGAAA 1500
Db |||||
Qy 1454 GAGATGAAGAAAGCGAGCCTTAAGCCTTACAGCCTTCAATTTAGATCATCATTCAGAAA 1513
Db |||||
Qy 1501 GCAATTCACACACACAGGATCGGATCAGTATCTCCGAGTTTCAAGAAATATATTTATC 1560
Db |||||
Qy 1514 GCAATTCACACACAGGATCGGATCAGTATCTCCGAGTTTCAAGAAATATATTTATGTT 1573
Db |||||
Qy 1561 CTTAAGATACAGAGGAGCTCTTTTAACTTAACCTTAGAGATAGCAGTCTTCAACAT 1620
Db |||||
Qy 1574 CTTAAGATACAGAGGAGCTCTTTTAACTTAACCTTAGAGATAGCAGTCTTCAACAT 1633
Db |||||
Qy 1621 AAAGAGAACTTGAATGGAATCTTATAGGACCATCTTAAAGTGGCCAAATGTA 1680
Db |||||
Qy 1634 AAAGAGAACTTGAATGGAATCTGATTTGGACCATCTTAAAGTGGCCAAATGTA 1693
Db |||||
Qy 1681 AATCTAAGAACTTAAAGATGAAAGTACAGGTTTGTACGAAGACTACTTTATTTT 1740
Db |||||
Qy 1694 AATCTAAGAACTTAAAGATGAAAGTACAGGTTTGTACGAAGACTACTTTATTTT 1753
Db |||||
Qy 1741 TACAGCCGAGAGTAAATTAATGCCAACTGTGATCTGGAATTTTGCAGGCGCAACAG 1800
Db |||||
Qy 1754 TACAGCCGAGAGTAAATTAATGCCAACTGTGATCTGGAATTTTGCAGGCGCAACAG 1813
Db |||||
Qy 1801 CTACAGGTTGTCGAGTTTACAGAAATTTCTTCTGAACTCTGAGAGGATGGGCA 1860
Db |||||
Qy 1814 CTACAGGTTGTCGAGTTTACAGAAATTTCTGCTGAGTCTGAGAGGATGGGCA 1873
Db |||||
Qy 1861 GGCTACTTAGAAGATCTAGTAAAGGATTTGTCAGTGGCTCAATGCTTCACTGGAATG 1920
Db |||||
Qy 1874 GGATACTTAGAAGATCTAGTAAAGGATTTGTCAGTGGCTCAATGCTTCACTGGAATG 1933
Db |||||
Qy 1921 AAACCCGAAAGAGTCTTCAAAATTAATGTTTATGACACCTTTAGTCAACACTACTTT 1980
Db |||||
Qy 1934 AAACCCGAAAGAGTCTTCAAAATTAATGTTTATGACACCTTTAGTCAACACTACTTT 1993
Db |||||
Qy 1981 TTATTTATGGAACACTTTCTTGCACCTCATGAGTAAATGCTGGAAGGATGAGT 2040
Db |||||
Qy 1994 TTATTTATGGAACACTTTCTTGTCTCATGAGTCAAAATGCTGGAAGGATGAGT 2053
Db |||||
Qy 2041 GTATTTCAAGTCTCTCTTAACTTTGCTCTTGAAGGATCACTTGTGTAAGCTT 2100
Db |||||
Qy 2054 GTCTTTCAAGTCTCTCTTAACTTTGCTCTTGAAGGATCACTTGTGTAAGCTT 2113
Db |||||
Qy 2101 ACTGTTTCAAGTCTCTCTTAACTTTGCTCTTGAAGGATCACTTGTGTAAGCTT 2160
Db |||||
Qy 2114 ACTGTTTCAAGTCTCTCTTAACTTTGCTCTTGAAGGATCACTTGTGTAAGCTT 2173
Db |||||
Qy 2161 TTAATGTCAGTCTGATGCTGAGTCTTCAACAACTTAAAGGATGATTTG 2220
Db |||||
Qy 2174 CTACGCGCAGAGTCTTGTGAGTCTGATGATTTTAAAGGATGATTTG 2233
Db |||||
Qy 2221 AGAGCTAATGTAATTTCTTAACTTTGCTCTTGAAGGATCACTTGTGTAAGCTT 2280
Db |||||
Qy 2234 AGAGCTAATGTAATTTCTTAACTTTGCTCTTGAAGGATCACTTGTGTAAGCTT 2293
Db |||||
Qy 2281 GATAAAAAAAGAGTCTTCTGAGTCTTGTGATATCTCTGATGAGGATGTAAGAC 2340
Db |||||
Qy 2294 GATAAAAAAAGAGTCTTCTGAGTCTTGTGATATCTCTGATGAGGATGTAAGAC 2353
Db |||||
Qy 2341 AAGGCGAATCTTCTGATGCTTCTTCAATGAGTCAACAGGTTTATCCACCTTGGAGAG 2400
Db |||||
Qy 2354 AAGGCGAATCTTCTGATGCTTCTTCAATGAGTCAACAGGTTTATCCACCTTGGAGAG 2413
Db |||||
Qy 2401 GGTTTGTCTCTGCTGAGATTTCTCTCAATTTCAAAAGGATTTTCTCTATCTGAATGAA 2460
Db |||||

```
Db      2414  GCGTCTCTACTGCTGAGATTTCTTCCATTCGGAAGGATTTCTTACTCTGAATGAA 2473
Qy      2461  AGAGGTTATGTAGCAAAACAAATGGAAAGTGCGCACAGGAA 2502
Db      2474  AGGGGTTATGTTGCAAAACAGTTGGAAAGTGCGCACAGGTA 2515

RESULT 5
LOCUS   AY417033
DEFINITION Homo sapiens HCM6089 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY417033
VERSION   AY417033.1 GI:39772993
KEYWORDS GSS.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1475)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
          Science 302 (5652), 1960-1963 (2003)
          14671302
REFERENCE 2 (bases 1 to 1475)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Direct Submission
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
          them based on alignment
          Location/Qualifiers
            source          1..1475
                           /organism="Homo sapiens"
                           /mol_type="genomic DNA"
                           /db_xref="taxon:9606"
            gene            <1..>1475
                           /locus_tag="HCM6089"

ORIGIN
Query Match      20.4%; Score 1047.6; DB 9; Length 1475;
Best Local Similarity 72.5%; Pred. No. 5.9e-269;
Matches 1066; Conservative 0; Mismatches 398; Indels 6; Gaps 1;

Qy      98  AACCTTCTGATAACTTAAGAGAGATTTCCAAAATGTGGCCAGATTGCGAGGAGTATCAA 157
Db      1  AACCTTCTGATAACTTAAGAGAGATTTCCAAAATGTGGCCAGATTGCGAGGAGTATCAA 60

Qy      158  ATATGAGAAAGCTAGGCCATCTGAATACTTAACTAGCTTCTTGTGATATTGGCCACA 217
Db      61  ATATGAGAAAGCTAGGCCATCTGAATACTTAACTAGCTTCTTGTGATATTGGCCACA 120

Qy      218  GTGAAGAAAAAAGCTGGGCTTTTCACTATGAGGATATCAATAATTTGTTGCGGTAGCTTTAT 277
Db      121  GTGAAGAAAAAAGCTGGGCTTTTCACTATGAGGATATCAATAATTTGTTGCGGTAGCTTTAT 180

Qy      278  TAAATGAAGCAAAAGAGATGCGAGCAGCAGGCTACAGAGGCTTCGATATCTCATCCAAG 337
Db      181  TAAATGAAGCAAAAGAGATGCGAGCAGCAGGCTACAGAGGCTTCGATATCTCATCCAAG 240

Qy      338  ACTCCAGTATTTCTCCAGAGAGTGTCTAAATTTGAAGTGGACTATTTAATAGCTAGGTGCA 397
Db      241  ACTCCAGTATTTCTCCAGAGAGTGTCTAAATTTGAAGTGGACTATTTAATAGCTAGGTGCA 300

Qy      398  TTGACATACACAGACGACGAGGTAGAGGACACAAAGCACTTTCGATTAGTCAGAAAGA 457

Db      301  TTGACATACACAGACGACGAGGTAGAGGACACAAAGCACTTTCGATTAGTCAGAAAGA 360
Qy      458  TGATTACTGTGAAATGCTTCTCTGTTCTCTAGTCTGTGACCAACTCAATTAATTCAGTTG 517
Db      361  TGATTACTGTGAAATGCTTCTCTGTTCTCTAGTCTGTGACCAACTCAATTAATTCAGTTG 420
Qy      518  GAAATGATGGACTTCAAGAAAGACAGAAATGGTCCGAGCATGCAATTCGCCATTAATCTGTG 577
Db      421  GAAATGATGGACTTCAAGAAAGACAGAAATGGTCCGAGCATGCAATTCGCCATTAATTCGCC 480
Qy      578  AACTAGCACATTCAAGATCCAGAGGTGTGGCCCTTTCAGAGGAGCTAAACACCAATATGA 637
Db      481  NNNNNNNNNNNNGAATCCAGAGGTGTGGCCCTTTCGAGGTGGACTAAACACCAATCTTGA 540
Qy      638  AAAATGTGATTTGATGGCAATTAAGTTCGAAATAAATGAGGCCCTAAATTAATTAATTTTGC 697
Db      541  AAAATGTGATCGATGGCAATTAAGTTCGAAATAAATGAGGCCCTAAATTAATTAATTTTGC 600
Qy      698  ACCTCTTTAAATCATCCAAAGACTCGGCAGTATGTGCGAGCTGTATGATAATTAGAGAA 757
Db      601  ACCTCTTTAAATCATCCAAAGACTCGACAGTATGTGCGAGCTGTATGATAATTAGAGNNN 660
Qy      758  TTTTAGCACCCCTATACTGATTTTCACTACACATAGTCCAGATACAGCTGAAGGACAGC 817
Db      661  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
Qy      818  TCAAGAAGACAGACAGACGACGATTTCTAGCCAGTAAATGGGAATCATAGCAACATTC 877
Db      721  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
Qy      878  GATCATGGCGAGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 937
Db      781  NNNNNNNNNNNNGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
Qy      938  TAGGAGTATTTGCAATACCAATATGGAATAAAGCGAGGTCTACTTGAAGTCTTTATG 997
Db      841  TAGGAGTATTTGCAATACCAATATGGAATAAAGCGAGGTCTACTTGAAGTCTTTATG 894
Qy      998  ATATATTTTCGTTCTCTCTACTGTTGTGACTGAGGAGTTTCATAGAAGCACTACTCAGTG 1057
Db      895  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 954
Qy      1058  TAGATCCAGGAGGTTCAGACAGTGGAGGCTTTAGATGGCTTTTGGCAGCTGAGG 1117
Db      955  NNNNNNNNNNGGAGTTTCCAAGACAGTTGAGGCTTTTCAATGAGCTTTTGGCAGCTGAGG 1014
Qy      1118  CAAAACTATTTCTCTCTCATCGTGCCAGATCCAGGCCAGACCTCATGTATAATTATTGG 1177
Db      1015  CAAAACTATTTCTCTCTCATCGTGCCAGATCCAGNNNNNNNNNNNNNNNNNNNNNNNNNN 1074
Qy      1178  CACTGATACCTCTCTGCAATTTATTCGTAATGGACTTTTAGAGGGTCTAGTTGAAGTGATA 1237
Db      1075  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1134
Qy      1238  CAAAACAGTGATGATCATATCTCAGTTAGAGTACCATCTCTTTTAGAGAGCTTTTACATA 1297
Db      1135  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1194
Qy      1298  TGGCAAAACACAATTTCTTCTCATTCACATAGCCATCATTTTACACTGCTTGCACACCTAA 1357
Db      1195  NNGCAAAACACAATTTCTTCTCATTCACATAGCCATCATTTTACACTGCTTGCACACCTAA 1254
Qy      1358  TGAATATGCTGCATCTCTTTTGATATCCCAAGGAAAAGAGACTGCGAGCAGTGCAGCCT 1417
Db      1255  TGAATATGCTGCATCTCTTTTGATATCCCAAGGAAAAGAGACTGCGAGCAGTGCAGCCT 1314
Qy      1418  TGAATGTTTTAAAACGCTTCTCATGAAATCAAGAACAGAGGAGCTTAAGCTTATAGTCTTC 1477
Db      1315  TGAATGTTTTAAAACGCTTCTCATGAAATCAAGAACAGAGGAGCTTAAGCTTATAGTCTTC 1374
Qy      1478  ATTTAGACACCACTATTTTCAAGAAAGCAATTCAGAACACACACAGAAACGGGATCATCTCC 1537
```


Db 1375 ATTTAGACCATTATTTCAGAAAGCAATTGCAACACACAGAAACGGGATCAGTATCTCC 1434

Qy 1538 GAGTTCAGAAAGATATATTATCTCTTAAGG 1567

Db 1435 GAGTTCAGAAAGATATATTATCTCTTAAGG 1464

RESULT 6
AY417035 1453 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM6089 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY417035
VERSION AY417035.1 GI:39772995
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1453)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Fierriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1453)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Fierriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (15-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source 1..1453
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1453
/locus_tag="HCM6089"

gene

ORIGIN

Query Match 17.2%; Score 880.8; DB 9; Length 1453;
Best Local Similarity 67.2%; Pred. No. 3.2e-224;
Matches 988; Conservative 0; Mismatches 454; Indels 28; Gaps 3;

Qy 98 AACCTTCTGATACTTAAGACAGATTCTCCAAAATGTGGCCAGATTGCAGGAGTATCAA 157

Db 1 AACCTTCTGATACTTAAGGAGATTCTCCAAAATGTGGCCAAAATGCAAGAGTATCAA 60

Qy 158 ATATGAGAAAGCTAGGCCATCTGAATAACTTTACTAAGCTTCTTTGTGATATGGCCACA 217

Db 61 ATATGAGAAAGCTGGCCATCTGAATACTTCACTAAGCTTCTTTGGATATGGCCATA 120

Qy 218 GTGAAGAAAACCTGGGCTTTCATATGAGGATATCAATAATTTGTTGGGTTAGCTTTAT 277

Db 121 GTGAAGAAAACCTGGGCTTTAAATATGAAGACATCAATAATTTGTTGGCTTAGCTTTAT 180

Qy 278 TAAATGAAGCAAAAGAGTGGCAGCAGCGGCTACGAGCGCTTCGATATCTCATCCAAAG 337

Db 181 TAAATGAAGCAAAAGAGTAAAGACGCGGGTTGGAGCGGCTCGCTATCTCATCCAAAG 240

Qy 338 ACTCCAGTATTTCTCCAGAGGTGCTAAAATTTGAAAGTGGACTATTTAATAGCTAGGTGCA 397

Db 241 ATTCCAGTATTTCTCCAGAGGTGCTAAAATTTGAAAGTGGACTATTTAATAGCTAGGTGCA 300

Qy 398 TTGCATACACAGAGCAACGAGGTAGAGGACACAAAGCACTTCGATTTAGTCAGAAAGA 457

Db 301 TTGCATTTACGACAGACAAACGAGGTGGAGAGGACACAAAGCCCTTCGCTTAGTCAGAAAGA 360

Qy 458 TGATTACTGTGAATGCTTCCCTTGTCTAGTCTCTGTGACCAACTCATTAAATTCAGTTG 517

Db 361 TGATTACTGTGAATGCTTCCCTTGTCTAGTCTCTGTGACCAACTCATTAAATTCGGGTG 420

Qy 518 GAAATGATGGACTTTCAAGAAAGAGACAGAAATGGTCCGAGCATGCAATTCCTATTCGTG 577

Db 421 GAAATGATGGGCTTTCAGGAACGAGACAGAAATGGTCCGGCTTGCATCCCAT----- 472

Qy 578 AACTAGCACTTCAGAAATCCAGAGGTGGTGGCCCTTCGAGGAGGACTAAACACCATATTGA 637

Db 473 -----GAACCCAGAGGTGGTAGCCCTTCGAGGAGGACTAAACACCATACTGA 519

Qy 638 AAAATGTGATTGATTGCCAATTAAAGTGAATAAATAGGCCCTTAATTACTACAAATTTTGC 697

Db 520 AAAATGTGATTGATTGCCAATTAAAGTGAATAAATAGGCCCTCTCATTAACAATTTTGC 579

Qy 698 ACCTTCTTAATCATCCAAAGACTCGGCAGTATGTGCGAGCTGATGTAGAAATTAGAGAGAA 757

Db 580 ACCTTCTTAATCATCCAAAGACCCGCGAGTATGTGCGAGCTGATGTAGAAATTAGAGNNN 639

Qy 758 TTTTAGCACCTTATCTGATTTTTCATCAGACATAGTCCAGATACAGCTGAAGACAGC 817

Db 640 NNN 699

Qy 818 TCAAAGAAAGACAGAGAAAGACGAGTATTTCTAGCCAGTAAAAATGGGAATCATAGCAACATTCC 877

Db 700 NNN 759

Qy 878 GATCATGGCGAGGTATTATTAATTTATGAACTCGGAAATTTCTGGGATCAGTCTCTAA 937

Db 760 NNNNNNNNNNNNGTATTATTAATTTATGTAAGCTCGGAAATTTCTGGGATACAGTCTCTAA 819

Qy 938 TAGAGTACTTTGATACACAAATATGAAATAGGCGAGGTCTACTTGAAGTGTCTTAAG 997

Db 820 TTGGAGTACTTTGATACACGAAATATGGAATAAGG-----NNNNNNNNNNNNNNNNNN 873

Qy 998 ATATATTTCTGCTTCTCTACCTGCTGACTGAGGAGTTTCATAGAGCACTACTCAGTG 1057

Db 874 NNN 933

Qy 1058 TAGATCCAGGAGGTTCCAAAGACAGTTGGAGGCTTTTCAGATGGCTTTGTGGCAGCTGAGG 1117

Db 934 NNNNNNNNGAAGATTCCAGGACAGTTGGAGGCTTTCCGATGGCTTTGTGGCAGCTGAGG 993

Qy 1118 CAAAAACTATTCTTCTCATCGCCAGATCCAGGCCAGACCTCATGGAATAATTTATTTGG 1177

Db 994 CAAAAACTATTCTTCTCCCATCGTCTAGATCCAGNNNNNNNNNNNNNNNNNNNNNNNNNN 1053

Qy 1178 CACTGATACTCTCTGCATTTATTTCGTAATGAGACTTTTAGAGGGTCTAGTTGAAGTGATA 1237

Db 1054 NNN 1113

Qy 1238 CAAAACAGTGTATCATATCTCAGTTAGAGTACCATCTCTTTTAGGAGAGCTTTTACATA 1297

Db 1114 NNN 1173

Qy 1298 TGGCAACACAAATTTCTTCTCATTCATATGACCATCATTTACACTGCTTGGCCAAACCTTAA 1357

Db 1174 NNGCAACACCATCTCGCTCTCACTCGCATACCCATCTACACTGCTTGGCCAAACCTTCA 1233

Qy 1358 TGAATATGGCTGGATCTTTTGTATATCCCAAGGAAAGAGACTCGCGAGCAGTCAGGCT 1417

Db 1234 TGAACATGGCTGGCTCTTCGATATTCGAAGGAAAGAGGCT--CGAGCCAGTGCAGCCC 1292

Qy 1418 TGAACCTTTTAAAAACGCTTCCATGAATGAAGAAACGAGGACCTTAAGCCTTTATAGTCTTC 1477

Db 1293 TGAACCTTTTGAACCGCTTCCATGAGATGAAGAGCGAGGACCTTAAGCCCTTACAGCCTTC 1352

Qy 1478 ATTTAGACCACTATTATTCAGAAAGCAATTCGAAACACACACAGAAACGGGATCAGTATCTCC 1537

Db 1353 ATTTAGATCATCATTTTCAAGAAAGCAATCGCAACTCAACCAAGCGGATCAGTATCTTC 1412


```

QY 1538 GAGTTCAGAAAGATATATTATCCCTTAAGG 1567
Db 1413 GAGTTCAGAAAGATATATTATTTGTTCTTAAGG 1442

RESULT 7
BQ049342 872 bp mRNA linear EST 29-MAR-2002
AGENCY: 7077338 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5787163
DEFINITION 5', mRNA sequence.
ACCESSION BQ049342
VERSION BQ049342.1 GI:19808682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 872)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM12877 row: a column: 20
High quality sequence stop: 637.
Location/Qualifiers
1. .872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5787163"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 16.1% Score 824.8; DB 5; Length 872;
Best Local Similarity 98.2%; Pred. No. 2.9e-209;
Matches 855; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 1973 ACTACTTTTATTTATTTGGAACACTTTCTTCCACCTCATGAGTTAAATGCTGGAAA 2032
Db 1 ACTACTTTTATTTATTTGGAACACTTTCTTCCACCTCATGAGTTAAATGCTGGAAA 60
QY 2033 AATGCAGTGTATTTTCAGTGTCTCTTAATCTTTGCTCTTGGAAACCAAGATCACTTGC 2092
Db 61 AATGCAGTGTATTTTCAGTGTCTCTTAATCTTTGCTCTTGGAAACCAAGATCACTTGC 120
QY 2093 TAAACCTTACTGTTCTTAGCTTGGACTATAGCAGATGATGGCTAGATGTCATCCTTT 2152
Db 121 TAAACCTTACTGTTCTTAGCTTGGACTATAGCAGATGATGGCTAGATGTCATCCTTT 180
QY 2153 CCAAAATTTTAACTGCGAGCTACTGATGCTGCGAGACTCTATGCAACAAATTAAGGG 2212
Db 181 CCAAAATTTTAACTGCGAGCTACTGATGCTGCGAGACTCTATGCAACAAATTAAGGG 240
QY 2213 TATTATTGAGAGCTAATGTTGAATTTCTTAATATTGGGAATTGAGTTGTTAGTGACCC 2272
Db 241 TATTATTGAGAGCTAATGTTGAATTTCTTAATATTGGGAATTGAGTTGTTAGTGACCC 300

```

```

QY 2273 AGCTACATGATAAAAAACAAACGATTTCTTCTGAAGCTCTTGATATATCCTCGATGAAGCAT 2332
Db 301 AGCTACATGATAAAAAACAAACGATTTCTTCTGAAGCTCTTGATATATCCTCGATGAAGCAT 360
QY 2333 GTGAAGAACAAGGCCAATCTTCTCATGCTCTCATTCAGATGAACACGAGCTTATCCACCTTG 2392
Db 361 GTGAAGAACAAGGCCAATCTTCTCATGCTCTCATTCAGATGAACACGAGCTTATCCACCTTG 420
QY 2393 GAGACACAGGGTTTCTTCTCTCTGCTGAGATTTCTTCTCCATTTCCAAAAAGGATTTTCTTATC 2452
Db 421 GAGACACAGGGTTTCTTCTCTCTGCTGAGATTTCTTCTCCATTTCCAAAAAGGATTTTCTTATC 480
QY 2453 TGAATGAAAGAGGTTATGTAGCAAAAAACAATTTGAAAAAGTGGCAACAGGGAATACAACTCCA 2512
Db 481 TGAATGAAAGAGGTTATGTAGCAAAAAACAATTTGAAAAAGTGGCAACAGGGAATACAACTCCA 540
QY 2513 AATATGTTGACTTGAATTTGAGGAACAACCTCAATGAGGCACTTACTTACCGGAAGCTG 2572
Db 541 AATATGTTGACTTGAATTTGAGGAACAACCTCAATGAGGCACTTACTTACCGGAAGCTG 600
QY 2573 TTGATGGTGATAACTATGTTCTGTCGAGTAACCAAAAGATTACAGCGTCTCTCACGCTTACC 2632
Db 601 TTGATGGTGATAACTATGTTCTGTCGAGTAACCAAAAGATTACAGCGTCTCTCACGCTTACC 660
QY 2633 TGCCTATACACCTTTATGGAACAACCTAGTACACCAATAAAACAGGCTGCCATTTGTTGGAAG 2692
Db 661 TGCCTATACACCTTTATGGAACAACCTAGTACACCAATAAAACAGGCTGCCATTTGTTGGAAG 720
QY 2693 TACAGAAATATTATTACAGAACTCTGTCGTAATGTTCTGTCACACCAAGATTGATAAGTGGG 2752
Db 721 TACAGAAATATTATTACAGAACTCTGTCGTAATGTTCTGTCACACCAAGATTGATAAGTGGG 780
QY 2753 AAGAAATTTAAAAAACTGAAAGCATCTCTTTGGGCTT--GGGAAATATCGCT--CATCAA 2809
Db 781 AAGAAATTTAAAAAACTGAAAGCATCTCTTTGGGCTTGGGAAATATCGCTTTCATCAA 840
QY 2810 ATTGGGGTCTCAATTTGCTACAGGAAGAAAA 2840
Db 841 TTGGGGGTCTCAATTTGCTACCGGAAGAAA 871

RESULT 8
AK075662 1977 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
DEFINITION enriched library, clone:1110055K10 product:unknown EST, full insert
sequence.
ACCESSION AK075662
VERSION AK075662.1 GI:26344462
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

```

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1977)
Adachi, J., Aizawa, K., Hanamura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Kanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saigo, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

COMMENT

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

Location/Qualifiers
1. .1977
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1110055K10"
/db_xref="taxon:10090"
/clone="1110055K10"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18-day embryo"
<1. .1089
/note="unnamed protein product; putative
unknown EST (GB|AV014646, evidence: BLASTN, 99%,
match=698)"
/codon_start=1
/protein_id="BAC35982.1"
/db_xref="GI:26344463"
/translation="SEALASPAKDVLTFTITMKANSFESRLTPSRFMKALSVALDK
EDLLSPNHTLQKSSVRMSVSATVGGDDVIGLALPVDINDIPIOLKVPYFQSKH
VPPDDRGMFHDGAGLSSGAGLVNPSHLLRQMSLTETMNSVHSDASLFLST
EDTGLQHTDDNLCVCIELLGFPQSNQLSSCSHSDLDQIDPYSWCELLTHNPLEV
VPKFGISGSDGAGSASSTELLGLVKTIIPDTPMCRWLLREKVLVNVN
LSSSVTKETHELTPIKEKYPQTFDDICLYSEVSHLLSHCTFRLQCRRFIQLFDV
QFLQMHAEAEVLAIPPIQPIVDESAES"

CDS

Query Match 16.0%; Score 818.6; DB 3; Length 1977;
Best Local Similarity 85.18; Pred. No. 1.7e-207;
Matches 927; Conservative 0; Mismatches 159; Indels 3; Gaps 1;
Qy 4042 TCTGAAGCTTTGGCATCTCCAGCAAAAGATGTGCTATTACTGTATACCATCACCATTGAAG 4101
Db 1 TCGGAAGCTTTGGCGTCCGCGGCAAAAGATGTGCTTTTACCGACACCATCACCATTGAAG 60
Qy 4102 GCCAACAAGCTTTGAGTCCAGATTAACCAACCAAGAGGTTTCATGAAGCTTTAAGTTATGCA 4161
Db 61 GCCAACAAGCTTTGAGTCCAGATTAACCAACCAAGAGGTTTATGAAGCTTTAAGTTATGCG 120
Qy 4162 TCATTAGATAAAGAGAGATTTATTGAGTCCCTATTATCAAAATACCTCGCAACGATCTTCC 4221
Db 121 TCATTAGATAAAGAGAGATTTATTGAGTCCCTATTATCAAAATACCTCGCAACGATCTTCC 180
Qy 4222 TCAGTGGCGTCCATGGTGTCCAGTGCACATATATGGGGGTTTCAGATGATTACATTTGGTCTT 4281
Db 181 TCAGTGGCGTCCATGGTGTCCAGTGCACATATATGGGGGTTTCAGATGATTACATTTGGTCTT 240
Qy 4282 GCTCTCCCGTGGATATAAATGATATATTCAGGTAAAGAGATATTCCTATTTTCAGACA 4341
Db 241 GCTCTCCCGTGGATATAAATGATATATTCAGGTAAAGAGATATTCCTATTTTCAGACA 300
Qy 4342 AAAAAACATACCAACCATGATGATCGAGTGCAGAGCATTTGGCCCATGATGCAGGAGGT 4401
Db 301 AAGCATGTGCTCCACAGATGACCGAGGTGCAGAGATGTTTTCCCATGATGGAGCAGGT 360
Qy 4402 CTTTCCATCTCGAACTGGAGGTCTTGTAAAAAATCTTTTTCATCTTGCTACGACGACGATG 4461
Db 361 CTTTCCATCTCGAACTGGAGGTCTTGTAAAAAATCTTTTTCATCTTGCTACGACGACGATG 420
Qy 4462 AGTCTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTGTTTGTAGAAAGATACA 4521
Db 421 AGTCTTACGGAATAATGAATTCATTCATTCAGATGCTCTCTGTTTGTAGAAAGATACA 480
Qy 4522 GAAGACATCGCATACAGGAACATACAGATGATAAATCGCTTTTATGTCGTCTGATTGAA 4581
Db 481 GAAGACATCGCATACAGGAACATACAGATGATAAATCGCTTTTATGTCGTCTGATTGAA 540
Qy 4582 ATTCTGGGTTTCCAGCCAGCAACCACTGAGTGCAATATGTAGTCAATTCAGACTTTCAA 4641
Db 541 CTCTGGGTTTCCAGCCAGCAACCACTGAGTGCAATATGTAGTCAATTCAGACTTTCAA 600
Qy 4642 GATATTCATATTCGTGATGGTGTGAGCAGACTATTCATTAATTCCTTTAGAGTGGTTCCT 4701
Db 601 GATATTCGTGATTCGTGATGGTGTGAGCAGACTATTCATTAATTCCTTTAGAGTGGTTCCT 660
Qy 4702 TCTAAGTTTTCGGGGATTTCTGGATGCAGTGCAGTGGGTGCTC--AAGAGGCTCAGCT 4758
Db 661 TCTAAGTTTTCGGGGATTTCTGGATGCAGTGCAGTGGGTGCTC--AAGAGGCTCAGCT 720
Qy 4759 AGCAGACCAAAAGACAGAAATTTGTTACTAGGTGTTTAAAAAATTCAGATGATACACCA 4818
Db 721 AGCAGACCAAAAGACAGAAATTTGTTACTAGGTGTTTAAAAAATTCAGATGATACACCA 780
Qy 4819 ATGTGGCGTATATCTCTTCGCAAAAGATTTCTAAGATTAGTCAATTAATTTAGAGTAGTTCA 4878
Db 781 ATGTGGCGTATCTCTCTTCGTAAGGAAGTCTAAGATTAGTGTCAATTTGAGTAGTTCA 840
Qy 4879 GTTTCACATTAATGTCTATGAGACTGGGCTTTTAAACATTAAGAGAGATGCTCCCAACA 4938
Db 841 GTTTCACATTAATGTCTATGAGACTGGGCTTTTAAACATTAAGAGAGATGCTCCCAACA 900
Qy 4939 TTTGATGACATATGCTCTTTACTCTGAGGTTTCCCATTTGCTGTGCACACTGCACATTGAGA 4998
Db 901 TTTGATGACATATGCTCTTTATTCTGAGTTTCTCACTGTTGTGCACACTGCACATTGAGA 960
Qy 4999 CTTCCGCTGTCGAGGTTTCATCAAGAAATTTATTTCAAGATGTACAGTTTCTACAATGAT 5058
Db 961 CTTCCGCTGTCGAGGTTTCATCAAGAAATTTATTTCAAGATGTACAGTTTCTACAATGAT 1020


```
Db 385 CTAAGCTGAAGTGGGCTATTTAATAGTAGTGGTCATTCGATTCAGCAGCAACGAG 444
Qy 421 GTAGAGAGGACACAGCACTTCGATTAGTCAGAAGAGATGATTACTGTGAATGCTTCCTTG 480
Db 445 GTGAGAGGACACAGCCCTTCGCTTAGTCAGAAGAGATGATTACTGTGAATGCTTCCTTG 504
Qy 481 TTTCTAGTCTGTGACCAACTCAATTAATTGCACTTGAAGAAATGATGCACTTCAAGAAAGA 540
Db 505 TTTCTAGTCTGTGCGCAACTCAATTAATTGCGGTTCGAAATGATGGCTTCAGGAACGA 564
Qy 541 GACAGATGTCGAGCATGCAATTCGCAATTCATCTGTGAATGCACTTCAAGATCCAGAG 600
Db 565 GACAGATGTCGCGGCTTCGATCGCCATCATCTGTGAATGCACTTCAAGATCCAGAG 624
Qy 601 GTGTGCGCCCTTCGAGGAGGACTAAACACCATATTTGAAATATGATTGTTGCAATTA 660
Db 625 GTGTGAGCCCTTCGAGGAGGACTAAACACCATATCTGAAATATGATTGTTGCAATTA 684
Qy 661 AGTCGAATTAATGAGGCGCCCTAATTAATTAATTTGCACTTCTTAATCAATCCAAAGACT 720
Db 685 AGTCGAATTAATGAGGCTCTCAATTAATTTGCACTTCTTAATCAATCCAAAGACT 744
Qy 721 CGCAGTATGTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 745 CAGCAGTATGTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 804
Qy 781 CACTACAGACATAGTCAGATACAGCTGAAGGAGGAGCTCAAGAGACAGAGACGACGA 840
Db 805 CACTACAGACATAGTCAGATACAGCTGAAGGAGGAGCTCAAGAGACAGAGACGACGA 864
Qy 841 TTTCTAGCCAGTAAATGGAATCATAGCAACATTCGATCATCGGCGAGCTATTATTAAT 900
Db 865 TTTCTAGCCAGTAAATGGAATCATAGCAACATTCGATCATCGGCGAGCTATTATTAAT 924
Qy 901 TTAT 904
Db 925 ATAT 928
```

```
RESULT 10
CD653846
LOCUS
DEFINITION
  AGENCOURT 14556896 NIA Human H1 Embryonic Stem Cell cDNA Library
  (Long) Homo sapiens cDNA clone IMAGE:30425290 5', mRNA sequence.
ACCESSION
  CD653846
VERSION
  EST.
KEYWORDS
  CD653846.1 GI:31892651
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 878)
REFERENCE
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Daniela S. Gerhard, Ph.D.
  Office of Cancer Genomics
  National Cancer Institute / NIH
  Bldg. 31 Rm10A07 Bethesda, MD 20892
  Email: cgabs-i@mail.nih.gov
  Tissue Procurement: Irene Gintis and Mahendra Rao, NIA
  cDNA Library Preparation: Yulan Piao and Minoru KO
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Cloning Distribution: MGC c
  can be found through the I.M.A.G.E. Consortium/LNL at:
  http://image.lnl.gov
  Plate: NDAM509 row: e column: 11
  High quality sequence stop: 704.
  Location/Qualifiers
  1..878
  /organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30425290"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/notes="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from Wicell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FUT3, SSA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGAGCGCGCCCTTTT-3'] from 3.45 of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."
```

ORIGIN

```
Query Match 14.7%; Score 753.2; DB 6; Length 878;
Best Local Similarity 94.9%; Pred. No. 4.6e-190;
Matches 782; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

Qy 1161 CATGGATATATTTGCGCACTGATCTCTGCACTATTTCGTAATGAGCTTTAGAGGG 1220
Db 11 CATGGATATATTTGCGCACTGATCTCTGCACTATTTCGTAATGAGCTTTAGAGGG 70

Qy 1221 TCTAGTTGAAGTGATAACAAACAGTGATGATCATATCTCAGTTAGAGCTACCATCCTTTT 1280
Db 71 TCTAGTTGAAGTGATAACAAACAGTGATGATCATATCTCAGTTAGAGCTACCATCCTTTT 130

Qy 1281 AGGAGAGCTTTTACATATATGCAAAACAAATTTCTTCCTCATTCATAGCCATCATTTACA 1340
Db 131 AGGAGAGCTTTTACATATATGCAAAACAAATTTCTTCCTCATTCATAGCCATCATTTACA 190

Qy 1341 CTGCTTGCACACCCCTAATGATGATGCTGCATCCTTTGATATCCCGAAGAAAGAGACT 1400
Db 191 CTGCTTGCACACCCCTAATGATGATGCTGCATCCTTTGATATCCCGAAGAAAGAGACT 250

Qy 1401 GCGAGCCAGTGCGAGCTTGAACCTGTTTAAACCCCTTCATGAATGAAGAAACGAGGACC 1460
Db 251 GCGAGCCAGTGCGAGCTTGAACCTGTTTAAACCCCTTCATGAATGAAGAAACGAGGACC 310

Qy 1461 TAAGCCTTATAGTCTTCATTAGACACCAATTTTCAGAAAGCAATTTGCAACACACAGAA 1520
Db 311 TAAGCCTTATAGTCTTCATTAGACACCAATTTTCAGAAAGCAATTTGCAACACACAGAA 370

Qy 1521 ACGGATCAGTATCTCGAGTTTCAGAAAGATATATTATCTTAAAGGATACAGAGGAGC 1580
Db 371 ACGGATCAGTATCTCGAGTTTCAGAAAGATATATTATCTTAAAGGATACAGAGGAGC 430
```


EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 867)
Woo, J.-T., Asai, Y., Zhu, M., Moates, J.M. and Magnuson, M.A.
Identification of thirty-six genes whose expression is markedly
affected by beta cell differentiation and de-differentiation
Unpublished (2001)
JOURNAL
COMMENT
Contact: Magnuson, M.A.
Department of Molecular Physiology and Biophysics
Vanderbilt University
747 LH, 21st Ave at Garland Ave. Nashville, TN 37232-0615, USA
Fax: 615 322 7236
Email: mark.magnuson@mcmail.vanderbilt.edu
PCR Primers
FORWARD: M13 forward
BACKWARD: M13 reverse
Insert Length: 867 Std Error: 0.00
Seq primer: T7, M13 reverse, custom
High quality sequence stop: 867
POLYA-No.

FEATURES
Location/Qualifiers
1..867
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2"
/db_xref="taxon:10090"
/clone="GKP2-237"
/cell_type="pancreatic endocrine tumor"
/lab_host="TOP10 (Invitrogen, Carlsbad, CA)"
/clone_lib="GKP2 subtraction cDNA library"
/note="Organ: Pancreas; Vector: pCR 2.1-TOPO (Invitrogen, Carlsbad, CA); Site: 1: EcoRI; Site: 2: BamHI; mRNA from GKP2 and GKP4 cells was used to perform subtraction suppression hybridization. Genes that are more highly expressed in GKP2 cells (an insulinoma derived cell), compared to GKP4 cells (a precursor cell line) were cloned pCR 2.1-TOPO and transformed into TOP10 competent bacteria."

ORIGIN
Query Match 13.9%; Score 710.2; DB 4; Length 867;
Best Local Similarity 88.7%; Pred. No. 1.6e-178;
Matches 769; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1763 ATGCCAACCTGGATCTGGATTTTGCACGCAACACAGCTACGGTTTGTAGTGGCCAGT 1822
DB 867 ACCTAGTCTGGATCTGGACTTGGCCAGCTTCCACGAGCTCACAGTTGTGGTTGTCACT 808
QY 1823 TTACAGAAATTTCTTGTGAATCTGAGAGGATGGCGAGGCTACTTAGAAGATCTAGTAA 1882
DB 807 TTACAGAAATTTCTGCTGAGTCTGAGAGGATGGCGAGGATCTTAGAAGATCTGCTGA 748
QY 1883 AGGATATTTGTTAGTGGCTCAATGCTTCTCATCTGGAATGAAACCCGAAAGAGTCTTCAAA 1942
DB 747 AAGATATTTGTTAGTGGCTCAATGCTTCTCATCTGGAATGAAACCCGAGAGAGCTTCA 688
QY 1943 ATAAATGGTTATTTGACCAACCTTAGTCAACACTACTTTTATTTATTTAGGAACACTTTCT 2002
DB 687 ATAAATGGTTATTTGACCAACCTTAGTCAACACTACTTTTATTTATTTAGGAACACTTTCT 628
QY 2003 GCCACCTCATGGAGTTAAATGCTGGAAATATGAGTGTATTTACAGTGTCTCTTAATC 2062
DB 627 GTCATCTCATGGAGTCAAAATGCTGGAAATATGAGTGTCTCTTAATC 568
QY 2063 TTTGCTCTTTGAAACCAAGATCACTTCTGTAACACTTACTTCTTACTAGCTTGGACTATA 2122
DB 567 TTTGCTCTTTGAAACCAAGATCACTTCTGTAACACTTACTTCTTACTAGCTTGGACTATA 508
QY 2123 GCAGAGATGGATGGCTAGAGTCATCTTTTCCAAAATTTTAACTGCAGCTACTGATGCTT 2182

Db 507 GCAGAGATGGATGGCTAGAGTCATCTTTTCCAAAATCTCTCACGGCAGCCACAGATGCTT 448
QY 2183 GCAGAGCTCTATGCAACCAAAACATTTTAAAGGGTATTTATGAGAGCTATGTTGAATTTCTTTA 224
DB 447 GCAGAGCTGTATGCAACCAAAACATTTTAAAGGGTATTTATGAGAGCTATGTTGAATTTCTTTA 388
QY 2243 ATAAATGGGGAAATTTGAGTTGTTAGTGCACCCAGCTACATGATAAAACCAAAACGATTTCTCT 230
DB 387 ACAATTTGGGGAAATTTGAGTTTACTAGTAACTCAGCTACATGATAAAACCAAAACGATTTCTCT 328
QY 2303 CTGAAGCTCTTGTATATCTCTCGATGAAGCATGTGAAGAACAAGGCCAATCTTATGCTCTCA 2366
DB 327 CTGAAGCGCTGGATATTTCTTGTATGAAGCTTGTGAAGAACAAGGCCAATCTTATGCTCTCTTA 268
QY 2363 TTCAGATGAAACCCAGCGTTTATCCACCTTTGGAGACAAGGGTTTGTCTCTCTGCTGAGAT 2422
DB 267 TTCAGATGAAGCAGCGCTTATCCACCTTTGGAGACAAGGGCTTGTCTTCTACTGCTGAGAT 208
QY 2423 TTCTCTCCATTTCCAAAAGGATTTTCTATCTGAATGAAAGAGGTTATGTAGCAAAACAAAT 248
DB 207 TTCTTTCCATTTCCGAAAGGATTTTCTTACCTGAATGAAAGGGGTTATGTGTCGAAACAGT 148
QY 2483 TGGAAAAGTGGCACAGGGAATACAATCCAAATATGTTGAATTGAGTGAAGGAAACAATCTCA 254
DB 147 TGGAAAAGTGGCACAGGGAATATAATTTCAAAGTATGTTGACTTGAATGAGGAAACAATCTTA 88
QY 2543 ATGAAGCAGCTTACTTACTTACCGGAAAGCCTGTTGATGTTGATGATAAATCTATGTTCTCGGAGTA 260
DB 87 ATGAAGCCTCTACTACTACATACAGGAAGCCTATTGATGGCGGATAAATACGTTGCTGCGAGTA 28
QY 2603 ACCAAAGATTACAGCGTCTCTCACGTCT 2629
DB 27 ATCAAGAGCTACAGCGACCTCATGCT 1

RESULT 13
BU746120/c
LOCUS
DEFINITION
Canis familiaris CDNA clone CH3#002_A09 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.
Expressed sequence tags from Canine heart
Unpublished (2003)
Other ESTs: CH3#002_A09T7
Contact: George AL
Division of Genetic Medicine
Vanderbilt University
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
Tel: 615 936 2660
Fax: 615 936 2661
Email: al.george@vanderbilt.edu
Insert Length: 1644 Std Error: 0.00
Seq primer: T3: ATTAACCTCACTAAAGGA
High quality sequence start: 61
High quality sequence stop: 871.
Location/Qualifiers
1..893
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CH3#002_A09"
/issue_type="heart"
/cell_type="mixed developmental stages (adult, 30 day - 40 day)"
/dev_stage="mixed developmental stages (adult, 30 day - 40 day)"

```
day fetal)"
/clone lib="Canine heart normalized cDNA Library in
pBluescript"
/Note="Organ: heart; Vector: pBluescript; Site 1: 5' of
vector NotI; Site 2: 3' of vector EcoRI; Tissue source:
dog heart (adult, 30 day - 40 day fetal), right and left
atria and ventricle. Dog breed - mixed (beagle, German
shepherd, pointer, Irish setter). Library construction:
oligo-dT primed"

ORIGIN
Query Match 13.7%; Score 702.6; DB 5; Length 893;
Best Local Similarity 92.0%; Pred. No. 1.7e-176;
Matches 763; Conservative 0; Mismatches 64; Indels 2; Gaps 2;

QY 2520 TGACTTGTAGTGGAGCAACTCAATGAAGCACTTACTTACCGAGCCCTGTTGATGG 2579
DB 861 TGACTTGTAGTGGAGCAACTCAATGAAGCACTTACTTACCGAGCCCTGTTGATGG 803

QY 2580 TGATAACTATGTTCTGTCGGAGT-AAACCAAGATTACAGCGTCTCACGTCCTACCTGCCTA 2638
DB 802 TGATAACTATGTTCTGTCGGAGTAAACCAAGATTACAGCGTCTCTCATGTCCTACCTGCCTG 743

QY 2639 TACACCTTTATGGACAACTAGTACACCAATAAACAGGCTGCCATTTGTTGGAAGTACAGA 2698
DB 742 TACACCTTTATGGACAACTAGTACACCAATAAACAGGCTGCCATTTGTTGGAAGTACAGA 683

QY 2699 ATATTATTACAACTCTGCTGTAATGTTGCTACACCAAGATTGATAGTGGGAAGAA 2758
DB 682 ATATTATTACCAACTCTGCTGTAATGTTGCTACACCAAGATTGATAGTGGGAAGAA 623

QY 2759 TTAATAAACTGAAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAAAATGGGGTC 2818
DB 622 TTAATAAACTGAAAGCATCTCTTTGGGCTTGGGAAATATCGGCTCATCAAAATGGGGTC 563

QY 2819 TCAATTTGCTACAGGAAGAAACGTTGATTCACAGATATATAAACTTGCACCAACAGTGTG 2878
DB 562 TCAATTTGCTACAGGAAGAAATGTTGATTCACAGATATATAAACTTGCACCAACAGTGTG 503

QY 2879 AAGTTCTTTCCATCAGAGGACCTGTGTATATGTTACTTGGGCTCATAGTAAACCAAC 2938
DB 502 AAGTTCTTTCCATCAGAGGACCTGTGTATATGTTACTTGGGCTCATAGTAAACCAAC 443

QY 2939 AAGGCTGTGATATCTTAAATGTCACACTGGGATGCTGTGAGGATAGTCACCAACATC 2998
DB 442 AAGGCTGTGATATCTTAAATGTCACACTGGGATGCTGTGAGGATAGTCACCAACATC 383

QY 2999 TGTGGCCAGTGTTCAGATGATGTGGAAACAACTCTGTAAATGAATCTTCACTATCCCAA 3058
DB 382 TGTGGCCAGTGTTCAGATGATGTGGAAACAACTCTGTAAATGAATCTTCACTATCCCAA 323

QY 3059 GCACCTTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTG 3118
DB 322 GCACCTTAAGTTTGAACCTCGGAGTCAACAGCTCTAGACATAATAGTGAAGTGAATCTG 263

QY 3119 TGCCATCGAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACAT 3178
DB 262 CACCATCAAGTATGTTTCATATTTGGAGGATGACCGGTTTGGCAGCAGCTCTACTAGTACAT 203

QY 3179 TTTTCCCTTGATATCAATGAAGTACAGAGCCAAATTTTATGACCGATCTGGACCCATAA 3238
DB 202 TTTTCCCTTGATATCAATGAAGTACAGAGCCAAATTTTATGACCGATCTGGACCCATAA 143

QY 3239 AGGATAAAATTCATCTCCCTTCTTCTTCTTCTAGTAAACTTGTGAAGATGTGATCTTAA 3298
DB 142 AGGATAAAATTCATCTCCCTTCTTCTTCTTCTAGTAAACTTGTGAAGATGTGATCTTAA 83

QY 3299 ATTGCTTACTTTCCTTAACAAAAACATCTGATGAGCAGTATCAACAA 3347
DB 82 ATTGCTTACTTTCCTTAAATAAAGAAATTCGATATCAAGCTTATCGATA 34

RESULT 14
```

BI601569

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI601569

603250782F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5302244 5',

mRNA sequence.

BI601569

BI601569.1 GI:15494508

EST.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 831)

NIH-MGC http://mgi.mcg.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9abps-remail.nih.gov

Tissue Procurement: Miklos Palakovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11765 row: d column: 21

High quality sequence stop: 771.

Location/Qualifiers

1..831

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5302244"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/clone_lib="NIH_MGC_96"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.3 kb and normalized to 10⁵. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity

Matches 742; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 1 ATGGGGCGGATCGGGCGGGCGGCTCTCTGAAGAACCTCCGAGTACGAGGGCGGAATGAC 60

DB 28 ATGGGGCGGATCGGGCGGGCGGCTCTCTGAAGAACCTCCGAGTACGAGGGCGGAATGAC 87

QY 61 AGCGCGGAGGAGAACGCTCCCGCTGATCTGACCGGAGAACCTCTGATAACTTAAGAGAG 120

DB 88 AGCGCGGAGGAGAACGCTCCCGCTGATCTGACCGGAGAACCTCTGATAACTTAAGAGAG 147

QY 121 ATCTCCAAAATGTGGCCAGATTGCGAGGAGTATCAATATGAGAAAGCTAGGCCATCTG 180

DB 148 ATCTCCAAAATGTGGCCAGATTGCGAGGAGTATCAATATGAGAAAGCTAGGCCATCTG 207

QY 181 AATACTTTTACTAGCTCTTTTGTGATATGCGCCACAGTGAAGAAAACCTGGCTTTTAC 240

DB 208 AATACTTTTACTAGCTCTTTTGTGATATGCGCCACAGTGAAGAAAACCTGGCTTTTAC 267

QY 241 TATGAGGATATCATAATTTTGTGCGGTGTAGCTTTTATAAATGAAGCAAAAGAGTGGCA 300

DB 268 TATGAGGATATCATAATTTTGTGCGGTGTAGCTTTTATAAATGAAGCAAAAGAGTGGCA 327

QY 301 GCAG-CAGGGCTACGAGCGCTTCGATATCTCATCCAAAGACTCCAGTATTTCTCCAGAGGT 359

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2005, 05:42:11 ; Search time 101 Seconds
(without alignments)
8659.715 Million cell updates/sec

Title: US-10-782-244-3
Perfect score: 8811
Sequence: 1 MAAIGRSLKSLNLRVRGRND.....AEAVLTPPKQPIVDTSAES 1708

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: +
1: uniprot_sprot: +
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8811	100.0	1708	2 Q6R327	Q6R327 homo sapien
2	8465	96.1	1708	2 Q6Q106	Q6Q106 mus musculus
3	8458	96.0	1708	2 Q6R174	Q6R174 mus musculus
4	7139	81.0	1385	2 Q68D75	Q68D75 homo sapien
5	6632	75.3	1275	2 Q8NCM6	Q8NCM6 homo sapien
6	5442	61.8	1051	2 Q86U57	Q86U57 homo sapien
7	4294	48.7	883	2 Q69Z40	Q69Z40 mus musculus
8	4165	47.3	824	2 Q8CBP2	Q8CBP2 mus musculus
9	4013	45.5	796	2 Q8R5D5	Q8R5D5 mus musculus
10	3986	45.2	801	2 Q8N3A0	Q8N3A0 homo sapien
11	1755.5	19.9	362	2 Q8BPH9	Q8BPH9 mus musculus
12	1661.5	18.9	1936	2 Q9VWU6	Q9VWU6 drosophila
13	1654.5	18.8	1936	2 Q7YU91	Q7YU91 drosophila
14	1476	16.8	308	2 Q8C2V9	Q8C2V9 mus musculus
15	1249	14.2	265	2 Q8N6M7	Q8N6M7 homo sapien
16	1233	14.0	255	2 Q8D2K3	Q8D2K3 m mus muscu
17	1233	14.0	1148	2 Q77203	Q77203 dictyosteli
18	1231	14.0	1153	2 Q7PSX6	Q7PSX6 anopheles g
19	1222	13.9	1148	2 Q76NV6	Q76NV6 dictyosteli
20	1089.5	12.4	219	2 Q6PDL2	Q6PDL2 mus musculus
21	1039	11.8	1541	2 Q7SBV4	Q7SBV4 neurospora
22	989.5	11.2	1309	1 ST16	Q9743 schizosacch
23	949.5	10.8	216	2 Q8CB99	Q8CB99 mus musculus
24	898.5	10.2	1168	2 Q6C3B5	Q6C3B5 yarrowia li
25	758.5	8.6	1176	2 Q6BYE3	Q6BYE3 debaryomyce
26	720.5	8.2	1430	1 YES3	P40061 saccharomyc
27	712.5	8.1	1348	2 Q75D08	Q75D08 ashbya gos
28	690.5	7.8	677	2 Q7KPC0	Q7KPC0 dictyosteli
29	655	7.4	1439	2 Q6FQM9	Q6FQM9 candida gla
30	618	7.0	1318	2 Q6CMH7	Q6CMH7 kluyveromyc
31	365	4.1	724	2 Q9TVP6	Q9TVP6 caenorhabdi

32	244	2.8	880	2 Q9XV48	Q9XV48 caenorhabdi
33	218	2.5	2319	1 AKA6	Q13023 homo sapien
34	217	2.5	2271	2 Q99074	Q99074 staphylococ
35	217	2.5	2271	2 Q7A362	Q7A362 staphylococ
36	216	2.5	2283	2 Q8VQ99	Q8VQ99 staphylococ
37	216	2.5	2297	2 Q9HGK6	Q9HGK6 candida alb
38	215	2.4	3317	2 Q8EWP8	Q8EWP8 mycoplasma
39	214	2.4	2275	2 Q8NUJ3	Q8NUJ3 staphylococ
40	214	2.4	2275	2 Q6G620	Q6G620 staphylococ
41	212.5	2.4	1878	2 Q6CP24	Q6CP24 kluyveromyc
42	210.5	2.4	534	2 Q94317	Q94317 schizosacch
43	205.5	2.3	1848	2 Q6CNL6	Q6CNL6 kluyveromyc
44	205.5	2.3	2424	2 Q7RN73	Q7RN73 plasmodium
45	204.5	2.3	1979	1 TRIA_HUMAN	Q15643 homo sapien

ALIGNMENTS

RESULT 1

Q6R327 Q6R327 PRELIMINARY; PRT; 1708 AA.
AC Q6R327; 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DB Rapamycin insensitive companion of mTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15268862; DOI=10.1016/j.cub.2004.06.054;
RA Dos D., Ali S.M., Kim D.H., Guertin D.A., Latek R.R.,
RA Erdjument-Bromage H., Tempst P., Sabatini D.M.;
RT "Rictor, a Novel Binding Partner of mTOR, Defines a Rapamycin-
RT Insensitive and Raptor-Independent pathway that Regulates the
RT Cytoskeleton.";
RL Curr. Biol. 14:1296-1302(2004).
DR EMBL; AY515854; AAS79796.1; -.
DR InterPro; IPR008938; ARM.
SQ SEQUENCE 1708 AA; 192215 MW; DB7B1E4A45DAE2AB CRC64;

Query Match 100.0%; Score 8811; DB 2; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAIGRSLKSLNLRVRGRNDSGEENVPLDILTREP	SDNLRILQNVARLQGVSNMRKLGHL	60
Db	1	MAAIGRSLKSLNLRVRGRNDSGEENVPLDILTREP	SDNLRILQNVARLQGVSNMRKLGHL	60
Qy	61	NNFTKLLCDIGHSEKLGPHYEDIIICLRALLNEA	KEVRAAGLRALRYLIQDSSIIQKV	120
Db	61	NNFTKLLCDIGHSEKLGPHYEDIIICLRALLNEA	KEVRAAGLRALRYLIQDSSIIQKV	120
Qy	121	LKLKVDYLIARCDIQOSNEVERTQALRLVRKMT	VNNSLPSSVTNSLTAAGNDGQER	180
Db	121	LKLKVDYLIARCDIQOSNEVERTQALRLVRKMT	VNNSLPSSVTNSLTAAGNDGQER	180
Qy	181	DRMVRACIATICEALQNPVVALRGGLNTILKNV	IDCQLSRINEALITITILLHNPKT	240
Db	181	DRMVRACIATICEALQNPVVALRGGLNTILKNV	IDCQLSRINEALITITILLHNPKT	240
Qy	241	RQYVRADVLERILAPYTDPHYRSPDTAEGQLKE	DREARFLASKMGIIATFRSWAGIN	300
Db	241	RQYVRADVLERILAPYTDPHYRSPDTAEGQLKE	DREARFLASKMGIIATFRSWAGIN	300
Qy	301	LCKPGNSGIQSLGVLCIPNWEIRRGILEVLYDI	FRPLPVTVEEFTEALLSVDPGRFQD	360
Db	301	LCKPGNSGIQSLGVLCIPNWEIRRGILEVLYDI	FRPLPVTVEEFTEALLSVDPGRFQD	360
Qy	361	SWRLSDGFVAEAKTILPHRARSRPDLMDNYLAL	ILISAFIRNGILLEGIVITNSDDHIS	420

Db 361 SWRLSDGFVAEAKTILPHRARSRLDMDNYLALILSAFIRNGLEGLVEVITNSDDHIS 420
Qy 421 VRATILLGELLMANTILPHSHSHLHCLPTLMMAASFDIPKPKRLASALNCLKRFH 480
Db 421 VRATILLGELLMANTILPHSHSHLHCLPTLMMAASFDIPKPKRLASALNCLKRFH 480
Qy 481 ENKKGPKPYSHLDHIIKATAHOKEDQYLRVQKDFILKDTTEALLINLRDSQVLQH 540
Db 481 ENKKGPKPYSHLDHIIKATAHOKEDQYLRVQKDFILKDTTEALLINLRDSQVLQH 540
Qy 541 KENLENNWNLIGTILKPNVNLNRYKDEQLHRFVRRLLYFYKPSKLYANLDLDFAKAQ 600
Db 541 KENLENNWNLIGTILKPNVNLNRYKDEQLHRFVRRLLYFYKPSKLYANLDLDFAKAQ 600
Qy 601 LTVGCGQTFEFLSEEDQGVLELDVQDIVOMLNASSGMPERSLQNGLLTTLTSHYF 660
Db 601 LTVGCGQTFEFLSEEDQGVLELDVQDIVOMLNASSGMPERSLQNGLLTTLTSHYF 660
Qy 661 LFTGTLSCHPHGVKMLEKSVFQCLNLCSLKNQDHLKLTVSSLDYSRDLGLARVILSKI 720
Db 661 LFTGTLSCHPHGVKMLEKSVFQCLNLCSLKNQDHLKLTVSSLDYSRDLGLARVILSKI 720
Qy 721 LTAATDACRLYATKHLRVLLRANVEFPNNGIIEILLVQLHDKNKTISSEALDILDEACED 780
Db 721 LTAATDACRLYATKHLRVLLRANVEFPNNGIIEILLVQLHDKNKTISSEALDILDEACED 780
Qy 781 KANLHALIQWKPALSHLGDKGLLLLRFLSIPKGSYLNERYGVAKOLEKMHREYNSKYV 840
Db 781 KANLHALIQWKPALSHLGDKGLLLLRFLSIPKGSYLNERYGVAKOLEKMHREYNSKYV 840
Qy 841 DLIEEQLEALATYRKPVDGNDVYRNSQRLQRPVHLYPIHLYGQLVHHKTCCHLEVN 900
Db 841 DLIEEQLEALATYRKPVDGNDVYRNSQRLQRPVHLYPIHLYGQLVHHKTCCHLEVN 900
Qy 901 IITELCRNVTPDLKWEELKKLSIWMALGNIGSSNWGLNLOENVPDILKLAKQCE 960
Db 901 IITELCRNVTPDLKWEELKKLSIWMALGNIGSSNWGLNLOENVPDILKLAKQCE 960
Qy 961 VLSIRGTCVVVLGIKAKTGCCDILKCHNWDVVRSHKHLWPVVPDVEQLCNELSIPS 1020
Db 961 VLSIRGTCVVVLGIKAKTGCCDILKCHNWDVVRSHKHLWPVVPDVEQLCNELSIPS 1020
Qy 1021 TSLNSESTSRHNSSESVPSSMFILEDRFGSSSTSTFFLDINEDTEPTFFVDRSGPIK 1080
Db 1021 TSLNSESTSRHNSSESVPSSMFILEDRFGSSSTSTFFLDINEDTEPTFFVDRSGPIK 1080
Qy 1081 DKNSFPFFASKLKVNRLNSLTLPNKKHRSDDPKGKLSSESKTNSRRIRLTPEPSVD 1140
Db 1081 DKNSFPFFASKLKVNRLNSLTLPNKKHRSDDPKGKLSSESKTNSRRIRLTPEPSVD 1140
Qy 1141 FNHSDDFPISTVQKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNGFTEHRENTSR 1200
Db 1141 FNHSDDFPISTVQKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNGFTEHRENTSR 1200
Qy 1201 ERLVVESTSHMKIRQSNTDTTSGISMSSSPSRETGVGDATMTDCGSMSTVVS 1260
Db 1201 ERLVVESTSHMKIRQSNTDTTSGISMSSSPSRETGVGDATMTDCGSMSTVVS 1260
Qy 1261 TKTIKTSHTLTPQSNHLSLSKNSVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFS 1320
Db 1261 TKTIKTSHTLTPQSNHLSLSKNSVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFS 1320
Qy 1321 YTSSRDAPGYATLKLQOORHPSLSHSEALASPAKOVLTDTITMKANGFESRLTPSRF 1380
Db 1321 YTSSRDAPGYATLKLQOORHPSLSHSEALASPAKOVLTDTITMKANGFESRLTPSRF 1380
Qy 1381 MKALSYASLDKEDLLSPINQNTLQSSSVRSWMVSSATYGGSDDYIGLALPVDINDIFQVK 1440
Db 1381 MKALSYASLDKEDLLSPINQNTLQSSSVRSWMVSSATYGGSDDYIGLALPVDINDIFQVK 1440
Qy 1441 DIPYFQTKNIPPHDRGARAFADAGGLPSGTGLVKNSFHLRQMSLITMNSIHSDA 1500

Db 1441 DIPYFQTKNIPPHDRGARAFADAGGLPSGTGLVKNSFHLRQMSLITMNSIHSDA 1500
Qy 1501 SLFLESTEDTGLQEHHTDDNCLYCVCIETILGFPQSNQLSAICSHSDPDIDIPYSDWCEQTH 1560
Db 1501 SLFLESTEDTGLQEHHTDDNCLYCVCIETILGFPQSNQLSAICSHSDPDIDIPYSDWCEQTH 1560
Qy 1561 NPLEVVPSPFSGISGSDGVSOGSASSSTKSTELLGVKTIIPDDTPMCRILLRKEVLRIV 1620
Db 1561 NPLEVVPSPFSGISGSDGVSOGSASSSTKSTELLGVKTIIPDDTPMCRILLRKEVLRIV 1620
Qy 1621 INLSSSVSTKCHETGLTITIKERYPQTFDDICIYSEVSHLSHCTFRLPCRRFIOELFQDV 1680
Db 1621 INLSSSVSTKCHETGLTITIKERYPQTFDDICIYSEVSHLSHCTFRLPCRRFIOELFQDV 1680
Qy 1681 QFLQMHHEAEAVLATPPKQPIVDTSAES 1708
Db 1681 QFLQMHHEAEAVLATPPKQPIVDTSAES 1708

RESULT 2
Q6Q106 PRELIMINARY; PRT; 1708 AA.
ID Q6Q106 AC Q6Q106;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Pianissimo.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LAF1;
RA Shiota C., Woo J.T., Lindner J., Shelton K.D., Zhu M.Y.,
RL Magnuson M.A.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY540053; AAS46920.1; -
DR InterPro; IPR008938; ARM.
SQ SEQUENCE 1708 AA; 191567 MW; C09CEEBABF1F5A00 CRC64;

Query Match 96.1%; Score 8465; DB 2; Length 1708;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1631; Conservative 40; Mismatches 36; Indels 2; Gaps 2;

Qy 1 MAAIGRGSLSKXLRVGRNDSGEENVPLDLTREPSDNLREILQNVARIQGVSNMRKLGHL 60
Db 1 MAAIGRGSLSKXLRVGRNDSGEENVPLDLTREPSDNLREILQNVARIQGVSNMRKLGHL 60
Qy 61 NNFTKLLCDIGHSEBKLGFFHYEDIIICLRALLNEAKEVRAAGLRALRYLIQDSSILQKV 120
Db 61 NNFTKLLCDIGHSEBKLGFFHYEDIIICLRALLNEAKEVRAAGLRALRYLIQDSSILQKV 120
Qy 121 LKLKVDYLIARCIDIQSNNEVERTQALRVKMTITVNASLPSSVTSNLSIAGVNDGLOER 180
Db 121 LKLKVDYLIARCIDIQSNNEVERTQALRVKMTITVNASLPSSVTSNLSIAGVNDGLOER 180
Qy 181 DRMVACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALTTILHLNHPKT 240
Db 181 DRMVACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALTTILHLNHPKT 240
Qy 241 RQYVRADVELEIRILAPYTDTFYRHSPDTAEGOLKEDREARFLASKWGIIATFRSWAGIIN 300
Db 241 RQYVRADVELEIRILAPYTDTFYRHSPDTAEGOLKEDREARFLASKWGIIATFRSWAGIIN 300
Qy 301 LCKPNSGSIQSLIGVLGICIPNMEIRRGLEVLVDIIFRLPLPVVTEFIEALLSVDFGRFO 360
Db 301 LCKPNSGSIQSLIGVLGICIPNMEIRRGLEVLVDIIFRLPLPVVTEFIEALLSVDFGRFO 360
Qy 361 SWRLSDGFVAEAKTILPHRARSRLDMDNYLALILSAFIRNGLEGLVEVITNSDDHIS 420
Db 361 SWRLSDGFVAEAKTILPHRARSRLDMDNYLALILSAFIRNGLEGLVEVITNSDDHIS 420

QY 421 VRATILGELLHMANVILPHSHSHHLLHCLPTLMNMAASFDIPKEKRLRASAALNCLKRFH 480
 Db 421 VRATILGELLHMANVILPHSHSHHLLHCLPTLMNMAASFDIPKEKRLRASAALNCLNRFH 480
 QY 481 EMKRGKPKPSYSLHDHIIIOKAIATHOKRQDQYLRVQKDIIFILKDTTEALLINLRDSQVLQH 540
 Db 481 EMKRGKPKPSYSLHDHIIIOKAIATHOKRQDQYLRVQKDIIFILKDTTEALLINLRDSQVLQH 540
 QY 541 KENLEWNNLIGTILKWPNNVLRNYKQDLHRFVRRLLYFYKPSKLYANLDLDPAKAQ 600
 Db 541 KENLDWNNLIGTILKWPNNVLRNYKQDLHRFVRRLLYFYKPSKLYASLDLAKSKQ 600
 QY 601 LTVVGCQPTFELSEEDQGYLEDVKDIIQVWLNASSGMPKPSLQNGLLTILSOHYF 660
 Db 601 LTVVGCQPTFELSEEDQGYLEDVKDIIQVWLNASSGMPKPSLQNGLLTILSOHYF 660
 QY 661 LFIGTSLCHPGVKNLEKCSVQCLNLCSLKNODHLLKLTVSSLDYSRDLARVILSKI 720
 Db 661 LFIGTSLCHPGVKNLEKCSVQCLNLCSLKNODHLLKLTVSSLDYSRDLARVILSKI 720
 QY 721 LTAATDACLRYATKHLRVLRYANVEFFNNWGIELLVTLQHDKNKTTISSEALDILDEACED 780
 Db 721 LTAATDACLRYATKHLRVLRYANVEFFNNWGIELLVTLQHDKNKTTISSEALDILDEACED 780
 QY 781 KANLHALIOMKPALSHLGDKGLILLRLPLSTPKFSYLNERYGVAKOLEKWHREYNSKYV 840
 Db 781 KANLHALIOMKPALSHLGDKGLILLRLPLSTPKFSYLNERYGVAKOLEKWHREYNSKYV 840
 QY 841 DLIEQNEALATTYRKPDVGNVYRRSNORLQRPVYLPVHYQLVGHKTKGCHLLVQON 900
 Db 841 DLIEQNEALATTYRKPDVGNVYRRSNORLQRPVYLPVHYQLVGHKTKGCHLLVQON 900
 QY 901 IITELCNVRTPDLKWEIEIKKASLWALGNISNWNGLNQLQENVIPDLKLAKQCE 960
 Db 901 IITELCNVRTPDLKWEIEIKKASLWALGNISNWNGLNQLQENVIPDLKLAKQCE 960
 QY 961 VLSIRGTCVYVGLIATKQCGCDILKCHNDVAVRHSRKLHWPVVDVQELCNELSPIS 1020
 Db 961 VLSIRGTCVYVGLIATKQCGCDILKCHNDVAVRHSRKLHWPVVDVQELCNELSPIS 1020
 QY 1021 TSLNSSTSSRRHNSSESVPSSMFILEDRFGSSSTSTPLDINEDTEPTFYDRSGPIK 1080
 Db 1021 TSLNSSTSSRRHNSSESVPSSMFILEDRFGSSSTSTPLDINEDTEPTFYDRSGPIK 1080
 QY 1081 DKNSFPFFASSKLVKNRILNSLTLPNKKHRSDDPKGKGLSSSEKTSNRRIRTLTEPSVD 1140
 Db 1081 DKNSFPFFASSKLVKNRILNSLTLPNKKHRSDDPKGKGLSSSEKTSNRRIRTLTEPSVD 1140
 QY 1141 FNSHDDTPISTVQKTLQETSFMGNKHIEDTGTSPSIGENDLKFTNPGTENHRENTSR 1200
 Db 1141 FNSHDDTPISTVQKTLQETSFMGNKHIEDTGTSPSIGENDLKFTNPGTENHRENTSR 1200
 QY 1201 ERLVSSSTSHMKIRSQSFNTDTTSGISMSSSPSRETGVGDATTMTDDCGSMSTVVS 1260
 Db 1201 ERLVSSSTSHMKIRSQSFNTDTTSGISMSSSPSRETGVGDATTMTDDCGSMSTVVS 1260
 QY 1261 TKTKTSHYLTPOSNHLSLKSNSVSLVPPGSSHTLPRRAQSLKAPSIAITIKSLADCNFS 1320
 Db 1261 TKTKTSHYLTPOSNHLSLKSNSVSLVPPGSSHTLPRRAQSLKAPSIAITIKSLADCNFS 1320
 QY 1321 YTSRDAFGYATLKRLOQRMHPSLSHSEALASPAKQVLTDTITMKANSFESRLTPSRF 1380
 Db 1321 YTSRDAFGYATLKRLOQRMHPSLSHSEALASPAKQVLTDTITMKANSFESRLTPSRF 1380
 QY 1381 MKALSASLKDLEDLSPINQNTLQSSSVRSWSSATYGGSDDYIGLALPVDINDIPQVK 1440
 Db 1381 MKALSASLKDLEDLSPINQNTLQSSSVRSWSSATYGGSDDYIGLALPVDINDIPQVK 1440
 QY 1441 DIPYFQKNIIPPHDDRGARAFADAGCLPSGTGVLKNSPHLLRQOQMSLTEINNSIHSDA 1500
 Db 1441 DIPYFQKNIIPPHDDRGARAFADAGCLPSGTGVLKNSPHLLRQOQMSLTEINNSIHSDA 1500
 QY 1501 SLFLESTEDTGLQHTDDNCLYCVCIIBILGFQPSNQLSSICSHSDLOQDIPYSWCQETIH 1560

Db 1500 SLFLESTEDTGLQHTDDNCLYCVCIIBILGFQPSNQLSSICSHSDLOQDIPYSWCQETIH 1559
 QY 1561 NPLEWVPSKSGISGSDGVSO-BGSASSTKSTELLIGVKTIPDDTPMCRILLRKEVRL 1619
 Db 1560 NPLEWVPSKSGISGSDGVSO-BGSASSTKSTELLIGVKTIPDDTPMCRILLRKEVRL 1619
 QY 1620 VINLSSSVTKCHBTGLTIKEKYPQTFDDICLYSEVSHLLSHCTFRLLPCRRFIQELFQD 1679
 Db 1620 VINLSSSVTKCHBTGLTIKEKYPQTFDDICLYSEVSHLLSHCTFRLLPCRRFIQELFQD 1679
 QY 1680 VQFLQMHAEAVLATPPKQPIVDUTSAES 1708
 Db 1680 VQFLQMHAEAVLATPPKQPIVDUTSAES 1708
 RESULT 3
 Q6R174 PRELIMINARY; PRT: 1708 AA.
 ID Q6R174
 AC Q6R174;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE AV03.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H;
 RA Loewith R.J., Hall M.N.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY497009; AAR89074.1; -
 DR InterPro; IPR008938; ARM.
 SQ SEQUENCE 1708 AA; 191525 MW; 452258052348A5BE CRC64;
 Query Match 96.0%; Score 8458; DB 2; Length 1708;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1629; Conservative 42; Mismatches 36; Indels 2; Gaps 2;
 QY 1 MAAGRGRSLKLVKVRGRNDSGEENVPLDLTRPSPDNLREILQNVARLQGVSNRKLHGL 60
 Db 1 MAAGRGRSLKLVKVRGRNDSGEENVPLDLTRPSPDNLREILQNVARLQGVSNRKLHGL 60
 QY 61 NNFTKLLCDIGHSEKLGPHYEDIICLRALLNEAKEVRAAGLRALRYLIQDSSIIQKV 120
 Db 61 NNFTKLLCDIGHSEKLGPHYEDIICLRALLNEAKEVRAAGLRALRYLIQDSSIIQKV 120
 QY 121 LKLVYDYLIAARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNLSIAVNGDGLQER 180
 Db 121 LKLVYDYLIAARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNLSIAVNGDGLQER 180
 QY 181 DRWVRACIAIICELALQNPVVALRGGLNTILKNVDCQLSRINEALITILLHLNHPKT 240
 Db 181 DRWVQACIAIICELALQNPVVALRGGLNTILKNVDCQLSRINEALITILLHLNHPKT 240
 QY 241 RQVVRADVELERILAPVTDPHYHSPDTAEGQLKEDREARFLASKMGIATFRSWAGIIN 300
 Db 241 RQVVRADVELERILAPVTDPHYHSPDTAEGQLKEDREARFLASKMGIATFRSWAGIIN 300
 QY 301 LCKPGNSGIQSLIGVLCIPNWEIRRGLELVYDI FRLPLPVWTEEFTEALLSVDPGRFQD 360
 Db 301 LCKPGNSGIQSLIGVLCIPNWEIRRGLELVYDI FRLPLPVWTEEFTEALLSVDPGRFQD 360
 QY 361 SWRLSDGFVAABEAKTILPHRARSRPLMDNYLALISAFIRNGLLEGLVEVITNSDDHIS 420
 Db 361 SWRLSDGFVAABEAKTILPHRARSRPLMDNYLALISAFIRNGLLEGLVEVITNSDDHIS 420
 QY 421 VRATILGELLHMANVILPHSHSHHLLHCLPTLMNMAASFDIPKEKRLRASAALNCLKRFH 480
 Db 421 VRATILGELLHMANVILPHSHSHHLLHCLPTLMNMAASFDIPKEKRLRASAALNCLNRFH 480

481 EMKRGPKPYSLHLDHIIQKAIATHQKRDQYLRVQKDIIFILKDTBEALLINLRDSQVLQH 540
Db
481 EMKRGPKPYSLHLDHIIQKAIATHHKKRDQYLRVQKDIIFVLKDTBEALLINLRDSQVLQH 540
QY
541 KENLENNWNLIGTILKWPNNVLRNRYKDEQLHFRFVRRLIYFYKPSKLYANLDLDFAKAKO 600
Db
541 KENLDWMDNLIGTILKWPNNVLRNRYKDEQLHFRFVRRLIYFYKPSKLYASLDLDAKAKO 600
QY
601 LTVVGQCFTEFLESSEDCGGYLEDVLDVQIOWLNASSGMKPERSLQNNGLLTTLSQHYF 660
Db
601 LTVVGQCFTEFLESSEDCGGYLEDVLDVQIOWLNASSGMKPERSLQNNGLLTTLSQHYF 660
QY
661 LFTIGTILSCHPHGVKMLEKCSVFQCLNLCSLKNQDHLKLTIVSSLDYSRDGLARVILSKI 720
Db
661 LFTIGTILSCHPHGVKMLEKCSVFQCLNLCSLKNQDHLKLTIVSSLDYSRDGLARVILSKI 720
QY
721 LTAATDACLRYATKHLRVLLRANVEFNWNGIELLVQIOWLNASSGMKPERSLQNNGLLTTLSQHYF 780
Db
721 LTAATDACLRYATKHLRVLLRANVEFNWNGIELLVQIOWLNASSGMKPERSLQNNGLLTTLSQHYF 780
QY
781 KANLHALIOMKPAISHLGDKGLLLRLFLSIPKGFSYLNERGVAKQEKWHREYNKYV 840
Db
781 KANLHALIOMKPAISHLGDKGLLLRLFLSIPKGFSYLNERGVAKQEKWHREYNKYV 840
QY
841 DLIEEQNEALTYRKPVDGDNVRRSNQRLQRPVYLPVHLYGQIVHHTKGCHLLEVN 900
Db
841 DLIEEQNEALTYRKPVDGDNVRRSNQRLQRPVYLPVHLYGQIVHHTKGCHLLEVN 900
QY
901 IITELCNVETPDLDWEETIKKLKASLWALGNIGSSNWGLNLIQENNVIPDILKAKQCE 960
Db
901 IITELCNVETPDLDWEETIKKLKASLWALGNIGSSNWGLNLIQENNVIPDILKAKQCE 960
QY
961 VLSIRGTCTVVLGLIAKTQCGDILKCHNWDVVRHGRKHLWVVPVDDVQVLCNELSIPS 1020
Db
961 VLSIRGTCTVVLGLIAKTQCGDILKCHNWDVVRHGRKHLWVVPVDDVQVLCNELSIPS 1020
QY
1021 TSLNSESTSRHNSSESVPSSMFILEDDRFSGSSTSTFFLDINDTPTFFYDRSGPIK 1080
Db
1021 TSLNSESTSRHNSSESVPSSMFILEDDRFSGSSTSTFFLDINDTPTFFYDRSGPIK 1080
QY
1081 DKNSFPFAGSKLVKKNILNSLTPNKKHRSDDPKGKLSSESNTSNRRIRLTLPSPVD 1140
Db
1081 DKNSFPFAGSKLVKKNILNSLTPNKKHRSDDPKGKLSSESNTSNRRIRLTLPSPVD 1140
QY
1141 FNSHDDFTPISTVQKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNGCTENHRENTSR 1200
Db
1141 LNHSEDFTS-SSAQSLQLEPFSFVGNKHELDAGSTPSIGENDLKFPKSGFTEHRENTSR 1199
QY
1201 ERLVVEGSTSHMKIRSQSFNTDTTSGISMSSSPSRETGVGDATTMDTCGSMSTVVS 1260
Db
1200 ERLVVEGASSHIKIRSQSFNTDTTSGISMSSSPSRETGVGDATTMDTCGSLSTVVS 1259
QY
1261 TKTIKTSHYLTPOSNHLSKSNVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFS 1320
Db
1260 TKTIKTSHYLTPOSNHLSKSNVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFS 1319
QY
1321 YTSRDAFCYATLKRLOOQRHPSLSHSEALASPAKOVLTDTITMKANGFESRLTPSRF 1380
Db
1320 YTSRDAFCYATLKRLOOQRHPSLSHSEALASPAKOVLTDTITMKANGFESRLTPSRF 1379
QY
1381 MKALSYASLDKEDLLSPINQNTLQSSSVRSMVSSATYGGSDDYIGLALPVDINDIFQVK 1440
Db
1380 MKALSYASLDKEDLLSPINQNTLQSSSVRSMVSSATYGGSDDYIGLALPVDINDIFQVK 1439
QY
1441 DIPYQTKNIIPHDDRGARAFADAGGLPSGTGGIVKNSFHLRQOQSLTEIMNSIHSDA 1500
Db
1440 DIPYQSKHVPVPPDDRGARAFADAGGLPSGTGGIVKNSFHLRQOQSLTEIMNSVHSDA 1499
QY
1501 SLFLESTEDTGLQHTDDNCLYCVCIETILGFPNSQLSAICSHSDFDQDIPYSDCEQTH 1560
Db
1500 SLFLESTEDTGLQHTDDNCLYCVCIETILGFPNSQLSAICSHSDLDQDIPYSDCEQTH 1559
QY
1561 NFLEVVPKSFSGISGSDGVSO-EGSASSTKSTELLGKVTIPDPTFMCRIILRKEVRL 1619

1560 NFLEVVPKSFSGISGSDGVSO-EGSASSTKSTELLGKVTIPDPTFMCRIILRKEVRL 1619
QY
1620 VINLSSSVSTKCHETGLTITKEKYPQTFDDICLYSEVSHLLSHCTFRLCRRFQIQLFOD 1679
Db
1620 VINLSSSVSTKCHETGLTITKEKYPQTFDDICLYSEVSHLLSHCTFRLCRRFQIQLFOD 1679
QY
1680 VOFLQHEEAEAVLATPPKQPIVDTSAES 1708
Db
1680 VOFLQHEEAEAVLATPPKQPIVDTSAES 1708
QY
RESULT 4
Q68DT5 Q68DT5 PRELIMINARY; PRT: 1385 AA.
AC Q68DT5; Q68DT5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686B11164.
GN Name=DKFZp686B11164;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus endometrial;
RG The German cDNA Consortium;
RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiegmann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR749280; CAH18135.1; --
DR InterPro; IPR008938; ARM.
KW Hypothetical protein.
KW SEQUENCE 1385 AA; 155815 MW; DALE1E24773F4BA4 CRC64;
Query Match 81.0%; Score 7139; DB 2; Length 1385;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1381; Conservative 1; Mismatches 3; Indels 38; Gaps 1;
QY 286 MGIIATFWSWAGIINLCKPGNSGQSGLIGVICIPNMEIRRGLELVYDIFRLPLPVVTEE 345
Db 1 MGIIATFWSWAGIINLCKPGNSGQSGLIGVICIPNMEIRRGLELVYDIFRLPLPVVTEE 60
QY 346 FTEALLSVDPGRFQDSWRLSDFGVAEAAKTILPHRARSRPLDMNYLALILSAFIRNGLL 405
Db 61 FTEALLSVDPGRFQDSWRLSDFGVAEAAKTILPHRARSRPLDMNYLALILSAFIRNGLL 120
QY 406 EGLVEVITNSDDHISVRATILGELLHMANTILPHSHSHHLLCLPTLNNMAASFDIPKEK 465
Db 121 EGLVEVITNSDDHISVRATILGELLHMANTILPHSHSHHLLCLPTLNNMAASFDIPKEK 180
QY 466 RLRSAAALNCLKRFHEMKRGPKPYSLHLDHIIQKAIATHQKRDQYLRVQKDIIFILKDT 525
Db 181 RLRSAAALNCLKRFHEMKRGPKPYSLHLDHIIQKAIATHQKRDQYLRVQKDIIFILKDT 240
QY 526 EALLINLRDSQVLOHKNLEWNNLIGTILKWPNNVLRNRYKDEQLHFRFVRRLIYFYKPS 585
Db 241 EALLINLRDSQVLOHKNLEWNNLIGTILKWPNNVLRNRYKDEQLHFRFVRRLIYFYKPS 300
QY 586 KLYANLDLDFAKAKOITVVGCOFTFLESSEDCGGYLEDVLDVQIOWLNASSGMKPER 645
Db 301 KLYANLDLDFAKAKOITVVGCOFTFLESSEDCGGYLEDVLDVQIOWLNASSGMKPER 360
QY 646 LQNNGLLTTLSQHYFLIGTILSCHPHGVKMLEKCSVFQCLNLCSLKNQDHLKLTIVSSL 705
Db 361 LQNNGLLTTLSQHYFLIGTILSCHPHGVKMLEKCSVFQCLNLCSLKNQDHLKLTIVSSL 420
QY 706 DYSRDLARVILSKILTAAATDACLRYATKHLRVLLRANVEFNWNGIELLVQIOWLNASSGMKPER 765
Db 421 DYSRDLARVILSKILTAAATDACLRYATKHLRVLLRANVEFNWNGIELLVQIOWLNASSGMKPER 480

```
Qy 766 ISSEALDILDEACEDKANLHALIQMKPALSHLQKGLLLLLRFLSLPKGFSYLNERYGVA 825
Db 481 ISSEALDILDEACEDKANLHALIQMKPALSHLQKGLLLLLRFLSLPKGFSYLNERYGVA 540
Qy 826 KOLEKWHREYNSKYVDLIEEQLNEALTYRKPVDGDNVRSNORLQRPVYLPYHLYGQ 885
Db 541 KOLEKWHREYNSKYVDLIEEQLNEALTYRKPVDGDNVRSNORLQRPVYLPYHLYGQ 600
Qy 886 LVHHTKGCHLLEVNIIITELCRNVRTPLDQKWEIKKLKASLWALGNIGSSNWGLNLQOE 945
Db 601 LVHHTKGCHLLEVNIIITELCRNVRTPLDQKWEIKKLKASLWALGNIGSSNWGLNLQOE 660
Qy 946 ENVIPDILKAKQCEVLSIRGTCVYVGLIAKTKQGCDDILKCHNDVAVRSHKHLWPVP 1005
Db 661 ENVIPDILKAKQCEVLSIRGTCVYVGLIAKTKQGCDDILKCHNDVAVRSHKHLWPVP 720
Qy 1006 DDVEQLCNELSSIPSTLSNSESSTRHNSSESVPSMFLEDDRGSSSTSTFFLDIN 1065
Db 721 DDVEQLCNELSSIPSTLSNSESSTRHNSSESVPSMFLEDDRGSSSTSTFFLDIN 780
Qy 1066 EDTEFTFYDRSGPIKDKNSPFPFASSKLVKNRILNSLTLPNKHRSDDPKGGKLSSES 1125
Db 781 EDTEFTFYDRSGPIKDKNSPFPFASSKLVKNRILNSLTLPNKHRSDDPKGGKLSSES 840
Qy 1126 TSNRRIRLTPEPSVDFPNHSDFTPISTVQKTLQLETSMGNKHIEDTGSTPSIGENDLKF 1185
Db 841 TSNRRIRLTPEPSVDFPNHSDFTPISTVQKTLQLETSMGNKHIEDTGSTPSIGENDLKF 900
Qy 1186 TKNEGTENHRENTSRRLVVESSSTSHMKIRSQSFNTDTTSGISSMSSPSRRTVGDA 1245
Db 901 TKNEGTENHRENTSRRLVVESSSTSHMKIRSQSFNTDTTSGISSMSSPSRRTVGDA 960
Qy 1246 TTMDTDCGSMSTVSTTKITKSHYLTPOSNHLSSKNSVSLVPPGSSHTLPRAAQSLKA 1305
Db 961 TTMDTDCGSMSTVSTTKITKSHYLTPOSNHLSSKNSVSLVPPGSSHTLPRAAQSLKA 1020
Qy 1306 PSIAIKSLADCNPSYSSRDAGYATLKRLOQRMHPSLSHSEALAPAKDVLFTDTIT 1365
Db 1021 PSIAIKSLADCNPSYSSRDAGYATLKRLOQRMHPSLSHSEALAPAKDVLFTDTIT 1080
Qy 1366 MKANSFESRLTPSRFMKALSASLDKEDLLSPINQNTLORSSSVRSVMSSATYGGSDYI 1425
Db 1081 MKANSFESRLTPSRFMKALSASLDKEDLLSPINQNTLORSSSVRSVMSSATYGGSDYI 1140
Qy 1426 GLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGLPSGTGGLVGNFSHLRQ 1485
Db 1141 GLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGLPSGTGGLVGNFSHLRQ 1181
Qy 1486 QMSLTIEMNSIHSDASLFELESTDTGLQEHDDNCLYCVCIILGFPQSNQLSAICSHSD 1545
Db 1182 -----ESTEDTGLQEHDDNCLYCVCIILGFPQSNQLSAICSHSD 1222
Qy 1546 FQDIPYSDWCETIHNPLEVVPKFSIGSGSDGVSQEGSASTKSTELLGVKTIIPDDT 1605
Db 1223 FQDIPYSDWCETIHNPLEVVPKFSIGSGSDGVSQEGSASTKSTELLGVKTIIPDDT 1282
Qy 1606 PMCRILARKEVLRVLINLSSSVTKCHETGLLTIKERYPOTFDDICLYSEVSHLLSHCTF 1665
Db 1283 PMCRILARKEVLRVLINLSSSVTKCHETGLLTIKERYPOTFDDICLYSEVSHLLSHCTF 1342
Qy 1666 RLPCCRRFIQELFQDVQFLQHEEAEAVLATPPKQPIVDTSAES 1708
Db 1343 RLPCCRRFIQELFQDVQFLQHEEAEAVLATPPKQPIVDTSAES 1385
```

RESULT 5

```
Q8NCM6 ID Q8NCM6 PRELIMINARY; PRT; 1275 AA.
AC Q8NCM6:
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1999 protein (fragment).
```

```
GN Name=KIAA1999;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [(1)]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=22051173; PubMed=12056414;
RA Ohara O., Nagase T., Mitsui G., Kohga H., Kikuno R., Hiraoka S.,
RA Takahashi Y., Kitajima S., Saga Y., Koseki H.;
RT "Characterization of size-fractionated cDNA libraries generated by the
RL DNA Reg. 9:47-57(2002).
DR EMBL; AB082530; BAC02708.1; -.
DR InterPro; IPR008938; ARM.
DR NON_TER 1
SQ SEQUENCE 1275 AA; 143548 MW; D5593DCC57CD0101 CRC64;
```

```
Query Match 75.3%; Score 6632; DB 2; Length 1275;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 434 ANTILPHSHSHLHCLPTLNMMAASFDIPKEKRLRASAALNCLRFHEMKRGPYSLH 493
Db 1 ANTILPHSHSHLHCLPTLNMMAASFDIPKEKRLRASAALNCLRFHEMKRGPYSLH 60
Qy 494 LDHIIQKAIATHQKRDQYLRVQKIDIFILKDTTEALLINLSDQVLQKENLWNWNLIGT 553
Db 61 LDHIIQKAIATHQKRDQYLRVQKIDIFILKDTTEALLINLSDQVLQKENLWNWNLIGT 120
Qy 554 ILKMPNVNLNRYKDEQLHRFRVRLLYFYKPSKLYANLDDFAKAKQLTVVGCOFTBFL 613
Db 121 ILKMPNVNLNRYKDEQLHRFRVRLLYFYKPSKLYANLDDFAKAKQLTVVGCOFTBFL 180
Qy 614 ESEEDGGYLEDLVKDIIVQWLNASSGMKPSRLQNNGLLTLSQHYFLFTGLSCHPHGV 673
Db 181 ESEEDGGYLEDLVKDIIVQWLNASSGMKPSRLQNNGLLTLSQHYFLFTGLSCHPHGV 240
Qy 674 KMLEKCSVFCIALNLSLKNQDHLKLTVSLSYSDRGLARVILSKILTAATDACRYAT 733
Db 241 KMLEKCSVFCIALNLSLKNQDHLKLTVSLSYSDRGLARVILSKILTAATDACRYAT 300
Qy 734 KHLRVLLRANVFPFNWNGIELLVTLQHDKNKTISSSEALDILDEACEDKANLHALIQMKA 793
Db 301 KHLRVLLRANVFPFNWNGIELLVTLQHDKNKTISSSEALDILDEACEDKANLHALIQMKA 360
Qy 794 LSHLGDGKLLLLRFLSLPKGFSYLNERYGVAKOLEKWHREYNSKYVDLIEEQLNEALT 853
Db 361 LSHLGDGKLLLLRFLSLPKGFSYLNERYGVAKOLEKWHREYNSKYVDLIEEQLNEALT 420
Qy 854 YRKPVDGDNVRSNORLQRPVYLPYHLYGQLVHHTKGCHLLEVNIIITELCRNVRTPD 913
Db 421 YRKPVDGDNVRSNORLQRPVYLPYHLYGQLVHHTKGCHLLEVNIIITELCRNVRTPD 480
Qy 914 LDKWEETIKKLKASLWALGNIGSSNWGLNLQOEENVPDILKAKQCEVLSIRGTCVYVGL 973
Db 481 LDKWEETIKKLKASLWALGNIGSSNWGLNLQOEENVPDILKAKQCEVLSIRGTCVYVGL 540
Qy 974 LIAKTQGCDDILKCHNDVAVRSHKHLWPVPDDVEQLCNELSSIPSTLSNSESSTRH 1033
Db 541 LIAKTQGCDDILKCHNDVAVRSHKHLWPVPDDVEQLCNELSSIPSTLSNSESSTRH 600
Qy 1034 NSESESVPSMFLEDDRGSSSTSTFFLDINEDTEFTFYDRSGPIKDKNSPFPFASSKL 1093
Db 601 NSESESVPSMFLEDDRGSSSTSTFFLDINEDTEFTFYDRSGPIKDKNSPFPFASSKL 660
Qy 1094 VKNRIILNSLTLPNKHRSDDPKGGKLSSESSTSNRRIRLTPEPSVDFPNHSDFTPISTV 1153
Db 661 VKNRIILNSLTLPNKHRSDDPKGGKLSSESSTSNRRIRLTPEPSVDFPNHSDFTPISTV 720
Qy 1154 QKTLQLETSMGNKHIEDTGSTPSIGENDLKFKNFCNTENHRENTSRRLVVESSSTSHM 1213
```


Db	721	QKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNFGTENHRENTSRRLVVESTSHM	780
Qy	1214	KIRSOSENTDTTSGISMSSSPSRETGVDAITMDTDCGSMSTVSTKTKTSHYLTTPQ	1273
Db	781	KIRSOSENTDTTSGISMSSSPSRETGVDAITMDTDCGSMSTVSTKTKTSHYLTTPQ	840
Qy	1274	SNHLSLSKNSVSLVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSYTSRDAFGVATL	1333
Db	841	SNHLSLSKNSVSLVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSYTSRDAFGVATL	900
Qy	1334	KLQOQRMHPSLSHSEALASPAKVDLFTDTITMKANSFESRLTPSRFMKALSASLDKED	1393
Db	901	KLQOQRMHPSLSHSEALASPAKVDLFTDTITMKANSFESRLTPSRFMKALSASLDKED	960
Qy	1394	LLSPINQNTLQRSSSVRSVSSATYGGSDYIGLALPVDINDIFQVKDIPYFQTKNIPPH	1453
Db	961	LLSPINQNTLQRSSSVRSVSSATYGGSDYIGLALPVDINDIFQVKDIPYFQTKNIPPH	1020
Qy	1454	DRGARAFADAGGLPSGTGGLVKNFSHLLRQOMSLTEIMNSIHSDASLFELESTEDTGLQ	1513
Db	1021	DRGARAFADAGGLPSGTGGLVKNFSHLLRQOMSLTEIMNSIHSDASLFELESTEDTGLQ	1080
Qy	1514	EHTDDNCLVCYCEIILGFQPSNQLSAICSHSDFODIPYSDWCEQIHNPLEVVPKFSGI	1573
Db	1081	EHTDDNCLVCYCEIILGFQPSNQLSAICSHSDFODIPYSDWCEQIHNPLEVVPKFSGI	1140
Qy	1574	SGCSGDSVQEGSASSTKSTELLIGVKTIPDDTPMCKRILLRKEVLRVILNLSVSTKCHE	1633
Db	1141	SGCSGDSVQEGSASSTKSTELLIGVKTIPDDTPMCKRILLRKEVLRVILNLSVSTKCHE	1200
Qy	1634	TGLLTKEKYPQTFDDICLYSEVSHLLSCTFRLPCRRTIOELFQDVQFLQWHEAEAVL	1693
Db	1201	TGLLTKEKYPQTFDDICLYSEVSHLLSCTFRLPCRRTIOELFQDVQFLQWHEAEAVL	1260
Qy	1694	ATPPKQPIVDTSAES 1708	
Db	1261	ATPPKQPIVDTSAES 1275	

RESULT 6

ID	Q86UB7	PRELIMINARY;	PRT; 1051 AA.
AC	Q86UB7;		
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	MGC39830	protein (Fragment).	
GN	Name=MGC39830;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
EC	TISSUE=Lymph;		
RA	MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heih F.,		
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly P.H.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Griinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		

RT		"Generation and initial analysis of more than 15,000 full-length human	
RL		and mouse cDNA sequences";	
RN	[2]	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RP		SEQUENCE FROM N.A.	
RC		TISSUE=Lymph;	
RA		Strausberg R.;	
RL		Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.	
DR		EMBL; BC051729; AAH51729.1; -	
DR		InterPro; IPR008938; ARM.	
FT		NON TER 1	
SQ		SEQUENCE 1051 AA; 117457 MW; 4D495C6A02B0D03 CRC64;	
		Query Match 61.8%; Score 5442; DB 2; Length 1051;	
		Best Local Similarity 99.8%; Pred. No. 1.2e-282;	
		Matches 1049; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Qy	658	HYFLFIGTSLCHPHGVKMLEKCSVFQCLNLCSLKQDHLKLTVSSLDYSRDLARVIL	717
Db	1	HYFLFIGTSLCHPHGVKMLEKCSVFQCLNLCSLKQDHLKLTVSSLDYSRDLARVIL	60
Qy	718	SKILTAATDACRLVATKHLVLRANVEFFNNWGIELLVTLQHDKNKTIISSEALDILDEA	777
Db	61	SKILTAATDACRLVATKHLVLRANVEFFNNWGIELLVTLQHDKNKTIISSEALDILDEA	120
Qy	778	CEDKANLHALIQMKPALSHLGDGKGLLLRFLSPKGFSYLNERGYVAKOLEKWHREYNS	837
Db	121	CEDKANLHALIQMKPALSHLGDGKGLLLRFLSPKGFSYLNERGYVAKOLEKWHREYNF	180
Qy	838	KYVDLIEQALNEALTTTRKPDVGDYVRRSNORLORPHVLPILHYGLVHHKGTGCHLLE	897
Db	181	KYVDLIEQALNEALTTTRKPDVGDYVRRSNORLORPHVLPILHYGLVHHKGTGCHLLE	240
Qy	898	VQNIITELCRNVTPDLDKWEIKKALASLWALGNIGSSNGLNLQENVIPDILKAK	957
Db	241	VQNIITELCRNVTPDLDKWEIKKALASLWALGNIGSSNGLNLQENVIPDILKAK	300
Qy	958	QCEVLISRGTCVYVGLIATKQGCIDILKCNWDAVHRSHKHLWPVPPDDVEQICNELSS	1017
Db	301	QCEVLISRGTCVYVGLIATKQGCIDILKCNWDAVHRSHKHLWPVPPDDVEQICNELSS	360
Qy	1018	IPSTLSINSESTSSRHNSSESVSPSSMFILDDRFSSSTSTFFLDINEDTEPTFYDRSG	1077
Db	361	IPSTLSINSESTSSRHNSSESVSPSSMFILDDRFSSSTSTFFLDINEDTEPTFYDRSG	420
Qy	1078	PIKDKNSPFFPASSKLVKRILNSITLKNKGRSSDPKGGKLSSEKTSNRRITLTETP	1137
Db	421	PIKDKNSPFFPASSKLVKRILNSITLKNKGRSSDPKGGKLSSEKTSNRRITLTETP	480
Qy	1138	SVDFNHSDDFTPISTVQKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNFGTENHRE	1197
Db	481	SVDFNHSDDFTPISTVQKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNFGTENHRE	540
Qy	1198	TSRERLVVESTSTSHMKIRSQSFNTDTTTSIGSSMSSSPSRETGVDAITMDTDCGSMST	1257
Db	541	TSRERLVVESTSTSHMKIRSQSFNTDTTTSIGSSMSSSPSRETGVDAITMDTDCGSMST	600
Qy	1258	VVSTKTKITSHYLTTPQSNHLSLSKNSVSLVPPGSSHTLPRAQSLKAPSIATIKSLADC	1317
Db	601	VVSTKTKITSHYLTTPQSNHLSLSKNSVSLVPPGSSHTLPRAQSLKAPSIATIKSLADC	660
Qy	1318	NFSYTSRDAFGVATLKRLOQOQRMHPSLSHSEALASPAKVDLFTDTITMKANSFESRLTP	1377
Db	661	NFSYTSRDAFGVATLKRLOQOQRMHPSLSHSEALASPAKVDLFTDTITMKANSFESRLTP	720
Qy	1378	SRFMKALSASLDKEDLSPINQNTLQRSSSVRSVSSATYGGSDYIGLALPVDINDIF	1437
Db	721	SRFMKALSASLDKEDLSPINQNTLQRSSSVRSVSSATYGGSDYIGLALPVDINDIF	780
Qy	1438	QVKDIPYFQTKNIPPHDDRGAFAHADAGGLPSGTGGLVKNFSHLLRQOMSLTEIMNSIH	1497
Db	781	QVKDIPYFQTKNIPPHDDRGAFAHADAGGLPSGTGGLVKNFSHLLRQOMSLTEIMNSIH	840

```
Qy 1498 SDASLFLESTEDTGLQEHDDNCLYVCIEILGFQPSNQLSAICSHSDPDIPYSDWCEQ 1557
Db 841 SDASLFLESTEDTGLQEHDDNCLYVCIEILGFQPSNQLSAICSHSDPDIPYSDWCEQ 900
Qy 1558 TIHNPVPSKSGISGSDGVSGSASSTKSTELLGVKTIPTDDTPMCRILLRKEVL 1617
Db 901 TIHNPVPSKSGISGSDGVSGSASSTKSTELLGVKTIPTDDTPMCRILLRKEVL 960
Qy 1618 RLVINLSSSVSTKCHETGLTITKEKYPTQDDICLYSEVSHLLSHCTFRPCRFIOELF 1677
Db 961 RLVINLSSSVSTKCHETGLTITKEKYPTQDDICLYSEVSHLLSHCTFRPCRFIOELF 1020
Qy 1678 QDVQFLOHMEAEAVLATPPKQPIVDTSAES 1708
Db 1021 QDVQFLOHMEAEAVLATPPKQPIVDTSAES 1051

RESULT 7
Q69240 PRELIMINARY; PRT; 883 AA.
AC Q69240;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MKIA1999 protein (fragment).
OS Name=MKIAA1999;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adult thymus;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT cDNAs Identified by Screening of Sequences of 500 Mouse KIAA-Homologous
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res 11:205-218(2004).
DR EMBL; AKI73326; BAD32604.1; -.
DR InterPro; IPR008938; ARM.
FT NON TER 1
SQ SEQUENCE 883 AA; 98005 MW; 9CE032EB5F383PA7 CRC64;

Query Match 48.78; Score 4294; DB 2; Length 883;
Best Local Similarity 92.84; Pred. No. 2.7e-221;
Matches 820; Conservative 30; Mismatches 32; Indels 2; Gaps 2;

Qy 826 KQLEKWHREYNSKYVDLIEBQLNEALTYTKRPVGDGNYVRRSNORLQRPVHYPILHYGQ 885
Db 1 KQLEKWHREYNSKYVDLIEBQLNEALTYTKRPVGDGNYVRRSNORLQRPVHYPILHYGQ 60
Qy 886 LVHKTGCHLLEQVNIITELCRNVRTPDLKDWEIEKIKLKSALWALGNIGSSNGLNLQ 945
Db 61 LVHKTGCHLLEQVNIITELCHNVRTPDLKDWEIEDKIKLKSALWALGNIGSSNGLNLQ 120
Qy 946 ENVIPDILKAKQCEVLISRGTCVYVGLIATKQCGDILKCNWADVRHSRHLWPVP 1005
Db 121 ENVIPDILKAKQCEVLISRGTCVYVGLIATKQCGDILKCHSWDSVRHSRHLWPVP 180
Qy 1006 DDVEQLCNELSSIPSTLSLNSESTSRHNSSESVSPSSMFLEDDRFSGSSTSTFFLDIN 1065
Db 181 DDVEQLCNELSSVPSTLSLNSESTSRHNSSESAFSPFMLEDDRFSGSSTSTFFLDIN 240
Qy 1066 EDTEFTYDRSGPIKDKNSPPFPFASLKVNRILNLTLPNKKHRSDDPKGKGLSSEK 1125
Db 241 EDAEAPAFYDRPGPIKDKNSPPFPFGSSKLVNRILNLTLPNKKHRSDDPKGKGLSSENK 300
Qy 1126 TSNRRIRTLTPSVDFNHSDDFTPISTVQKTLQLETFSFMGNKHIEDTGSTPSPGENDLKF 1185
Db 301 TSNRRIRTLTPSVDFNHSDDFTS-SSAOKSLQLEPFSFVGNKHLEDAGSTPSPGENDLKF 359
```

RESULT 8

```
O8CBF2 PRELIMINARY; PRT; 824 AA.
ID Q8CBF2
AC Q8CBF2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:9630041c24 product:hypothetical ARM repeat
DE structure containing protein, full insert sequence. (Fragment).
GN Name=4921505C17rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
```

Mon Apr 11 08:54:04 2005

60,770 full-length cDNAs. ";
 Nature 420:563-573 (2002).
 [4]
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum; DOI=10.1101/gr.145100;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 [5]
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 [6]
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK036149; BAC29321.1; -
 DR MGD; MGI:1926007; 4921505C17rik.
 DR InterPro; IPR008938; ARM.
 KW Hypothetical protein.
 FT NON TER 824
 SQ SEQUENCE 824 AA; 93509 MW; C1B80EBB2F818F22 CRC64;
 Query Match 47.3%; Score 4165; DB 2; Length 824;
 Best Local Similarity 98.2%; Pred. No. 1.9e-214;
 Matches 809; Conservative 11; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 MAAIGRSLKLVNRGSDGSENVPLDITRPSNLRILQVAVLRQVSNRKLGLH 60
 Db 1 MAAIGRSLKLVNRGSDGSENVPLDITRPSNLRILQVAVLRQVSNRKLGLH 60
 Qy 61 NNFTKLCIDTGHSEKLGPHYEDIIICRLALNEAKEVAAGRLRYLIQDSSILQKV 120
 Db 61 NNFTKLCIDTGHSEKLGPHYEDIIICRLALNEAKEVAAGRLRYLIQDSSILQKV 120
 Qy 121 LKLVYDLIARCDIQSNEVERTQALRVKMTITVNASLPSPSVNLSIANGDGLQER 180
 Db 121 LKLVYDLIARCDIQSNEVERTQALRVKMTITVNASLPSPSVNLSIANGDGLQER 180
 Qy 181 DRWVRACIATCEALQNPVVALRGGLNTILKNVIDCOLSRNEALITLILHNPHT 240
 Db 181 DRWVRACIATCEALQNPVVALRGGLNTILKNVIDCOLSRNEALITLILHNPHT 240
 Qy 241 RQYVRADVELERILAPYTDHYRSPDTAEQKEDREARFLASKMGIIATFRSWAGIIN 300
 Db 241 RQYVRADVELERILAPYTDHYRSPDTAEQKEDREARFLASKMGIIATFRSWAGIIN 300
 Qy 301 LCKPGNSGIQSLIGVLCIPNMEIRRGLEVLVDIFRLPLVVTDFEIALLSVDPGRFQD 360

Db 301 LCKPGNSGIQSLIGVLCIPNMEIRRGLEVLVDIFRLPLVVTDFEIALLSVDPGRFQD 360
 Qy 361 SWRLSDGFVAEAAKTILPHRARSRLPDMONYLALILSAFIRNGLLEGLVEVITNSDDHIS 420
 Db 361 SWRLSDGFVAEAAKTILPHRARSRLPDMONYLALILSAFIRNGLLEGLVEVITNSDDHIS 420
 Qy 421 VRATILGELLHMANITLPHSHSHHLLHCLPPTLMNMAASFDIPKEKRLASAAALNCLGRFH 480
 Db 421 VRATILGELLHMANITLPHSHSHHLLHCLPPTLMNMAASFDIPKEKRLASAAALNCLGRFH 480
 Qy 481 EMKRGPKPYSLHLDHIIQKAIATHOKRQDOYLVRQKIDIFILKOTERALLINRDSOVLOH 540
 Db 481 EMKRGPKPYSLHLDHIIQKAIATHOKRQDOYLVRQKIDIFILKOTERALLINRDSOVLOH 540
 Qy 541 KENLEWNNWNLGITLKWPNVNLNRYKDEQLHFRVRLLYFYKPSKLYANLMDLDFAKAKQ 600
 Db 541 KENLDWNNWNLGITLKWPNVNLNRYKDEQLHFRVRLLYFYKPSKLYANLMDLDFAKAKQ 600
 Qy 601 LTVVGCOFTBFLLESEEDGGYLEDLVKDIVOVLNASSGMKPKPERSLQNNGLLTLSQHYF 660
 Db 601 LTVVGCOFTBFLLESEEDGGYLEDLVKDIVOVLNASSGMKPKPERSLQNNGLLTLSQHYF 660
 Qy 661 LFIOTLSCHPHGVKMLEKCSVFOCLNLCSIKNODHLLKLTVSLDYSRDGLARVILSKI 720
 Db 661 LFIOTLSCHPHGVKMLEKCSVFOCLNLCSIKNODHLLKLTVSLDYSRDGLARVILSKI 720
 Qy 721 LTAATDACRLVATKHLRVLLRANVEFFNNWGIELVTQLHDKNKTTISSEALDILDEACED 780
 Db 721 LTAATDACRLVATKHLRVLLRANVEFFNNWGIELVTQLHDKNKTTISSEALDILDEACED 780
 Qy 781 KANLHALIQMKPALSHLGDGLKLLLLRFLSLPKGFSYINRGYV 824
 Db 781 KANLHALIQMKPALSHLGDGLKLLLLRFLSLPKGFSYINRGYV 824
 RESULT 9
 Q8R5D5 PRELIMINARY; PRT; 796 AA.
 ID Q8R5D5; AC Q8R5D5; DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE 4921505C17rik protein (fragment).
 GN Name=4921505C17rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]

```
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC022906; AAH22906.1; -.
DR MGD; MGI:1926007; 4921505C17rik.
FT InterPro; IPR008938; ARM.
FT NON_TER 1
SQ SEQUENCE 796 AA; 91055 MW; C4FE28FCFD63F0B7 CRC64;

Query Match 45.5%; Score 4013; DB 2; Length 796;
Best Local Similarity 97.8%; Pred. No. 2.5e-206;
Matches 769; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

Qy 145 QALRLVRKMITVNASLPSSVNTSLIAVNGDGLQERDMVRACIAIICELALQNPVAL 204
Db 1 QALRLVRKMITVNASLPSSVNTSLIAVNGDGLQERDMVRACIAIICELALQNPVAL 60
Qy 205 RGLNTILKNVIDCOLSRINEALITITLHLNHPKTRQYVRADVELERILAPYTDHYRH 264
Db 61 RGLNTILKNVIDCOLSRINEALITITLHLNHPKTRQYVRADVELERILAPYTDHYRH 120
Qy 265 SPTAEGQLKDEARFLASKMGIATFRSWAGIINLCKPGNSGIOSLIGVLCIPNMEIR 324
Db 121 SPTAEGQLKDEARFLASKMGIATFRSWAGIINLCKPGNSGIOSLIGVLCIPNMEIR 180
Qy 325 RGLLEVLYDIIFRLPLPVVTEEFIEALISVDPGRFQDSWRLSDFVAEAKTILPHRARSR 384
Db 181 RGLLEVLYDIIFRLPLPVVTEEFIEALISVDPGRFQDSWRLSDFVAEAKTILPHRARSR 240
Qy 385 PDLMDNYLALILSAPIRNGLLGLEIVITNSDDHISVRATILGELLHMANTILPHSHSH 444
Db 241 PDLMDNYLALILSAPIRNGLLGLEIVITNSDDHISVRATILGELLHMANTILPHSHSH 300
Qy 445 HLHCLPTLMNMAASFDIPKEKRLRASAAALCLRFHEMKRGPKPSLHLDHIIQKAIAT 504
Db 301 HLHCLPTLMNMAASFDIPKEKRLRASAAALCLRFHEMKRGPKPSLHLDHIIQKAIAT 360
Qy 505 HOKRDQYLRVQKIDIFILKOTEEALLNLNRDSQVLQKHENLEWNNWLIIGTILKWPNNLNRN 564
Db 361 HOKRDQYLRVQKIDIFILKOTEEALLNLNRDSQVLQKHENLEWNNWLIIGTILKWPNNLNRN 420
Qy 565 YKDEQLHRFVRRLLYFYKPSKLYANLDLDFAKAQLTVVGCQFTBFLSEEDGQYLE 624
Db 421 YKDEQLHRFVRRLLYFYKPSKLYASLDLDAKSKQLTVVGCQFTBFLSEEDGQYLE 480
Qy 625 DLVKDQIVQWLNASSGKPKERSLQNGLLITLSOHYFLFIGTILSCHPHGVKMLEKCSVFQC 684
Db 481 DLVKDQIVQWLNASSGKPKERSLQNGLLITLSOHYFLFIGTILSCHPHGVKMLEKCSVFQC 540
Qy 685 LNLCSLKQNDHLLKLTVSLSLDYSRDGLARVILSKILTATDACLRYATKHLRVLLRANV 744
Db 541 LNLCSLKQNDHLLKLTVSLSLDYSRDGLARVILSKILTATDACLRYATKHLRVLLRANV 600
Qy 745 EFPNNNGIELLVTLQHDKNKTIISSEALDILDEACEDKANLHALIQMKPALSHLGDKGLLL 804
Db 601 EFPNNNGIELLVTLQHDKNKTIISSEALDILDEACEDKANLHALIQMKPALSHLGDKGLLL 660
Qy 805 LARFLSIPKGFSLNERNYGYAKOLEKWHREYNKYVDLIEQLNEALTTRYKRVGDGNVY 864
Db 661 LARFLSIPKGFSLNERNYGYAKOLEKWHREYNKYVDLIEQLNEALTTRYKRVGDGNVY 720
Qy 865 RRSNQLRQPHVYLPVHLVYQVHHKTGCHLLEVSQSIITELCRNVRTPDLDKWEEKKCLK 924
Db 721 RRSNQLRQPHVYLPVHLVYQVHHKTGCHLLEVSQSIITELCRNVRTPDLDKWEEKKCLK 780
Qy 925 ASLWAL 930
Db 781 ASLWAL 786

RESULT 10
QBN3AO
```

```

Db 661 SKFSGISGCGSDVQBSASSTKSTELLGKVTIPDDTPMCRILLRKEVLRLVNLSSV 720
QY 1628 STKCHETGLTTIKKYPQTFDDICLYSEVSHLSHCTFRLPCRFTQELFDQVQFLQME 1687
Db 721 STKCHETGLTTIKKYPQTFDDICLYSEVSHLSHCTFRLPCRFTQELFDQVQFLQME 780
QY 1688 EAEAVLATPPKQPIVDTSAES 1708
Db 781 EAEAVLATPPKQPIVDTSAES 801

RESULT 11
Q8BPH9 PRELIMINARY; PRT; 362 AA.
AC Q8BPH9
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:111005K10 product:unknown EST, full insert sequence.
DE (Fragment)
DE Name=4921505C17rik;
CN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]_
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075662; BAC35882.1; -.
DR MGD; MGI:1926007; 4921505C17rik.
FT NON_TER
SQ SEQUENCE 362 AA; 39953 MW; E3F4309B226F4545 CRC64;

Query Match 19.9%; Score 1755.5; DB 2; Length 362;
Best Local Similarity 93.4%; Pred. No. 5.5e-86;
Matches 338; Conservative 10; Mismatches 13; Indels 1; Gaps 1;

QY 1348 SEALASPAKDVLTDTITMKANSFESRLTPSRFMKALSYASLDKEDLLSPINQNTLQRSS 1407
Db 1 SEALASPAKDVLTDTITMKANSFESRLTPSRFMKALSYASLDKEDLLSPINQNTLQRSS 60

QY 1408 SVRSWSSATYGGSDDYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFANDAGG 1467
Db 61 SVRSWSSATYGGSDDYIGLALPVDINDIFQVKDIPYFQSKHVPDPDRGARFMSHDGAG 120

QY 1468 LPSGTGGVLVKNSFHLLRQMSLTETMNSIHSDASLFLESTEDTGLQEHDDNCLYCVCTE 1527
Db 121 LSSGAGGLVKNSFHLLRQMSLTETMNSVHSDASLFLESTEDTGLQEHDDNCLYCVCTE 180

QY 1528 ILGFQPSNQLSAICSHSDFQDIPYSDWCEQTIHNPVVPKFSIGSCDGVSO-EGSA 1586
Db 181 LLGFQPSNQLSSICSHSDDLQDIPYSDWCEQTIHNPVVPKFSIGSCDGVSOEGSA 240

QY 1587 SSTKSTELLGKVTIPDDTPMCRILLRKEVLRLVNLSSSVSTKCHETGLTTIKKYPQ 1646
Db 241 SSTKSTELLGKVTIPDDTPMCRILLRKEVLRLVNLSSSVSTKCHETGLTTIKKYPQ 300

QY 1647 FDDICLYSEVSHLSHCTFRLPCRFTQELFDQVQFLQMEAEAVLATPPKQPIVDTSA 1706
Db 301 FDDICLYSEVSHLSHCTFRLPCRFTQELFDQVQFLQMEAEAVLATPPKQPIVDTSA 360

QY 1707 ES 1708
Db 361 ES 362

RESULT 12
Q9VMJ6 PRELIMINARY; PRT; 1936 AA.
AC Q9VMJ6
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CG8002-PA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H., Blazej R.G., Champe M., Pfeiffer B.D.,

```

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Fogle C., Gabrieli A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorral J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22426065; PubMed=12537568;
 RX Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RL [3]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22426070; PubMed=12537573;
 RX Kaminiker J.S., Bergman C.M., Krommiller B., Carlson J., Swirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RL [4]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22426069; PubMed=12537572;
 RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RL [5]
 RP SEQUENCE FROM N.A.
 RP FlyBase.
 RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RP [6]
 RP SEQUENCE FROM N.A.

Db 892 NKRYVLLVEADTHSSITLHNRNEDGYYSRRNCNQRPTQVPPNVAHPHYLQMAQTQGMWA 951
Qy 896 LEVQNIITELCRNVRTPDLDKWEIKKAKASUWALGNIGSSNWNGLNLOEEN--VIPDIL 953
Db 952 LRKHGDLPLLELLRAKCTDDAECLKAAIWAHAHASTHNGIEYFVELNARLYEKLI 1011
Qy 954 KLAQCEVLISIRGTCVYVGLIAKTQKGDILKCHNDVARSRKHLPV-VPDVEQLC 1012
Db 1012 VLVTKCEVYVRATCFSGALGIAGTQAGANILFKLNLWSLVKDNKMPVHPHED----- 1066
Qy 1013 NELSSIPSTLSNLSSTSSRHNSSESVSPSMFILEDD---RFGSSSTSTFFLDINEDTE 1069
Db 1067 -----WMSQVTPVRH--YYEDVPADNMTWMDVIERFYETG-SDFWHMLADQI- 1112
Qy 1070 PTFYRSGIPKKNSPFFPASSKLVKNRILNLSLTPNKKHRSDDPKGKGLSEKTSNR 1129
Db 1113 -----DASG-----GGGVAGTVVG-----MHPVQDPNATITTTDS----- 1144
Qy 1130 RIRLTPEPSVDFNHSDDFTP-----ISTVQTKLOLETSFMGNKH---IEDTGSTPSIGEN 1181
Db 1145 -VRTTT-----DDVHPRMLVGAAKSKTLPESNLRQGHORSSEKTTDVIS-- 1192
Qy 1182 DLKFTKNFGTGNHRENTSRERLIVESTSSHKIRSQSFNTDTTSGISSMSSSPSRET 1241
Db 1193 --LLSGSVCTMSG-----AVLYPTPYQHRIRVNSC-TDSNTSGVSSCESVTGRTAA 1240
Qy 1242 GVDATTMDTDCGSMSTVSTKIKTHYLTPOSNHLSKNSVSLVPGSSHTLPRRAQ 1301
Db 1241 AYAAAANELQQPPLSPSPMSNLL-----ESDELFRNQLATSMLP----- 1281
Qy 1302 SLKAPSIATIKSLADCNFSYTSRDAFGYATLKRLOQOORMHPSLGHSEALAPAKDVLPT 1361
Db 1282 -----STLSPMA-----MKGTVQURSL-RKHSRPFVSESSAEFFDFABILD 1322
Qy 1362 DTITTKANSFESRLTPSRFMKALSASLDKE--DL---LSPINQNTLQSSSVRSVMSSA 1416
Db 1323 PEVQMR-----KLDWTHSHRRLKVRSLDRQLSDVYRRLSADBNVLPLTNAKPLPN 1376
Qy 1417 TVGGSDDYIGLALPVDINDI FQVKDIP-YFQTKNIPPHDRGARAPAHADAGLPGSTGGL 1475
Db 1377 DLKG-PCYAGICLPKNVLDFPTRLNLSRTYVSRDIQDDIVGI----- 1418
Qy 1476 VKNSFHLLQOMSLTBMNSIHSASLFLSESTEDTG-----LQE---HTDONCLYC 1523
Db 1419 --NLNTMLRPOCLNDSLNEGDESIVSSLSVSSASRRQTLQOQAKHSRSLCLHC 1476
Qy 1524 VCIEILGPOFQSNLSAICSHSD---FQDIP---YSDWCQETIHNPLEVVP--SKPSGIS 1574
Db 1477 -----ARGPRQORNDGSHNGGGGASLAPCELYSSAAAAALVAAGIQAGPVLAKKSGSS 1530
Qy 1575 GCSGDSVQSE-GSASSTKSTELLIGVKTIPDDTPMCRILLRKEVLRVLAVINLSSSVSTKCHE 1633
Db 1531 --AQGASALGADISPHSPESMLSESLPD-----RLTASLTVNQRLANPVSQKSK 1581
Qy 1634 TGLLTKEKYPTQFDICLYSEVSHLLSHCTFRLPCRPTQELFQDVQVF----- 1682
Db 1582 MALLELKQKHPHAFQIDICLYSEACKTIGRSSYEMIAARRFLQELFLDLNFDTSFVVEPQLII 1641
Qy 1683 -----LQHEEAEAVLAT-----PPKQPIV 1702
Db 1642 GARKFSAEKEAEATAPTAMQAMPFKQITL 1670

RESULT 13

Q7YU91
ID Q7YU91 PRELIMINARY; PR: 1936 AA.
AC Q7YU91;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE S16537p.
GN ORFNames=CG8002;
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreesnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungal C.J., Nunoo J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT009929; AA022398.1; --
DR FlyBase; FBgn0031006; CG8002.
DR InterPro; IPR008938; ARM.
SQ SEQUENCE 1936 AA; 214447 MW; 589325908AC70A1D CRC64;

Query Match 18.8%; Score 1654.5; DB 2; Length 1936;
Best Local Similarity 27.9%; Pred. No. 2e-79;
Matches 510; Conservative 315; Mismatches 699; Indels 305; Gaps 54;

Qy 8 RSLKNLRVRGNDSEENVPLDLTPESDNLREILQNVARLOGVNMKGLHNLFTKLL 67
Db 13 RSKLQRIKVSQDP-EDFYRLDPQPSAAENAFIYMLC-LEETDRDTRKFLNLALSLC 70
Qy 68 CDI--GHSEKLGHYEDIILCLRLALLNEAKEVEAAGLRALRYLIQDSSILQKVLKLV 125
Db 71 LGARKGSHSNIRFTTEBLLCYLSASLVHTTQVRAAALRTIRYALSTPNDIKTNAQL 130
Qy 126 DYLIARCDIQOSNEVERTQALRLVRKMTVNASLFPSSVTNSLIANGDGLQERDMVR 185
Db 131 QHLLCESIDLMLKNDDERVQALKVRKMLAIPEDISPAVVRCLVSLADSGEENDNLLR 190
Qy 186 ACIAIICELALONPEVALRGGLNTILKNVDQCLSRNEALITILHLNPKTRQYVR 245
Db 191 ACLATLABFAVLNPAIIVCGVTSITRNVLECHNPRIAESLCGVLLYLLLEWPTQNI-- 248
Qy 246 ADVELERILAPYTDPHYRHSPTASGOLKEDREARFLASKMGIATFRSWAGIINLC 305
Db 249 CGVRUDCLAAFCYDTYQS---IVDKKDAELRYTSCRALLSVLSRWTGTLFCFDP 305
Qy 306 N-SGQSLIGVLICFNMEIRGLLEVLIDYIFRLPLPVVTEFEIALSVDPGRFQDS 364
Db 306 KPSGLKAIVDALYLNQIEVRKAILDLYELLTLPQPTWTDYVAVALQAVDPDFQDTW 365
Qy 365 SDGFVAEAKTILPHRARSRLMDNYLALILSAPINRGLLEGLVEVITNSDDHISVR 424
Db 366 SNGFVSAEGRSILTLAARAPSVVQHLALMLYCFLETGLNALVEVVVSSQDFSVQ 425
Qy 425 ILLGELLHMANTILPHSHSHLHCLPTLMNMAASFDIPKEKRLASAAALNCLKRHE 484
Db 426 VILGKILQMLHLPDIDICCTSPALPTLVSHATL-----GNQANAAVAALQNY 479
Qy 485 RGPKPYSLHLDHIQ--KAIAATHQKRDQYLRVQKD---IFILKOTEE----- 526
Db 480 QRPASRSLFLDSIIQGGALIQRLFR-RHLNVQEQAGPVLQQLQETAEEAFAVPP 538
Qy 527 -----ALLINLR-----DSOVLQHK 541
Db 539 GTLDRSLDSVSSDESNSQASTSSRSPRLKRLKFLPQALYDNFRNFRNLLTDSR 598
Qy 542 ENLEWNWNLIGTILKWPVNVLNENYDEQLHRFVRRLLYFYKPSKLYANLDFAKAK 601
Db 599 DAHLMDWDVITILK---SNLIRKLDYTGKFKRLVDFYKPRNNRFSHODL--V 653
Qy 602 TV---VGCQFTFELLESE-DCQGYLEDLVKDI VQWLNA--SSGMKPERSLQNG 655
Db 654 PTYVSAGLDLIDVLLSSNELECMRFITDFSDISQLAATVTSNRAHDCFLSPQH 713
Qy 656 SQHYFLFICTLSCHPHGVKMLEKCSVFOCLNLCSLKNQDHLKLTVSSLSRDCGL 715
Db 714 CQOYFLIYIGRMCRKTVKGLIEVLKNTVFEYLINLVRVTHVCVVKLIVSGNISY 773


```
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Taya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK087858; BAC40027.1; -.
DR MGD; MGI:1926007; 4921505C17Rik.
DR InterPro; IPR008938; ARM.
KW Hypothetical protein.
SQ SEQUENCE 308 AA; 34738 MW; 41375330A3729D54 CRC64;

Query Match      16.8%; Score 1476; DB 2; Length 308;
Best Local Similarity 97.7%; Pred. No. 4e-71;
Matches 292; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAIGRSLKRLVRGRNDSGEENVPLDITREPSDNLREILQNVARLQGVSNMRKLGHL 60
DB 1 MAAIGRSLKRLVRGRNDSGEENVPLDITREPSDNLREILQNVARLQGVSNMRKLGHL 60
QY 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRLLALNEAKEVRAAGLRALRYLIQDSSILQKV 120
DB 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRLLALNEAKEVRAAGLRALRYLIQDSSILQKV 120
QY 121 LKLKVDYLITARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNSLIANGNDGLOER 180
DB 121 LKLKVDYLITARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNSLIANGNDGLOER 180
QY 121 LKLKVDYLITARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNSLIANGNDGLOER 180
DB 121 LKLKVDYLITARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNSLIANGNDGLOER 180
QY 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITTIHLHLNHPKT 240
DB 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITTIHLHLNHPKT 240
QY 241 RQYVRADVELERILAPYTFDHYRHSPTAEGQLKEDREARFLASKWGIIATFRSWAGII 299
DB 241 RQYVRADVELERILAPYTFDHYRHSPTAEGQLKEDREARFLASKWGIIATFRSWAGML 299

RESULT 15
Q8N6M7 PRELIMINARY; PRT; 265 AA.
AC Q8N6M7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MGC39830 protein.
GN Name=MGC39830;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC029608; AAH29608.1; -.
DR InterPro; IPR008938; ARM.
SQ SEQUENCE 265 AA; 29828 MW; 5634C7B650EA5A9C CRC64;

Query Match      14.2%; Score 1249; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 4.5e-59;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAIGRSLKRLVRGRNDSGEENVPLDITREPSDNLREILQNVARLQGVSNMRKLGHL 60
DB 1 MAAIGRSLKRLVRGRNDSGEENVPLDITREPSDNLREILQNVARLQGVSNMRKLGHL 60
QY 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRLLALNEAKEVRAAGLRALRYLIQDSSILQKV 120
DB 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRLLALNEAKEVRAAGLRALRYLIQDSSILQKV 120
QY 121 LKLKVDYLITARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNSLIANGNDGLOER 180
DB 121 LKLKVDYLITARCIDIQOSNEVERTQALRLVRKMITVNASLFPSSVTNSLIANGNDGLOER 180
QY 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITTIHLHLNHPKT 240
DB 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITTIHLHLNHPKT 240
QY 241 RQYVRADVELE 251
DB 241 RQYVRADVELE 251

Search completed: April 9, 2005, 15:14:47
Job time : 110 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2005, 13:24:55 ; Search time 27 Seconds
(without alignments)
6086.601 Million cell updates/sec

Title: US-10-782-244-3

Perfect score: 8811

Sequence: 1 MAALGRGRSLKLVGRND.....ABAVLATPPKQPIVDTSAES 1708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1225	13.9	1148	2 JC5984	CAMP-dependent cel
2	989.5	11.2	1309	2 T39379	sexual differentia
3	720.5	8.2	1430	2 S50596	hypothetical prote
4	365	4.1	724	2 T21539	hypothetical prote
5	244	2.8	880	2 T21538	hypothetical prote
6	217	2.5	2271	2 F90073	hypothetical prote
7	210.5	2.4	534	2 T39903	serine-rich protei
8	200.5	2.3	1131	2 T41144	hypothetical serin
9	200.5	2.3	2345	2 T13829	Tpr homolog - fru
10	197.5	2.2	3187	2 JC5837	364K Golgi complex
11	197	2.2	3507	2 T34513	hypothetical prote
12	196.5	2.2	1306	2 T25370	MSB2 protein - yea
13	192	2.2	1419	2 T30531	agglutinin-like ad
14	188.5	2.1	725	2 A41258	a-agglutinin core
15	188	2.1	824	2 T21675	hypothetical prote
16	185.5	2.1	1173	2 T43527	sp8 protein - fiss
17	185.5	2.1	1727	2 T50073	myosin-like coiled
18	184	2.1	1502	2 T48309	hypothetical prote
19	182.5	2.1	2211	1 KFB05	coagulation factor
20	182	2.1	2224	1 KFHU5	coagulation factor
21	180.5	2.0	1072	2 A86827	hypothetical prote
22	178.5	2.0	2493	2 T49303	probable membrane
23	178.5	2.0	2748	2 S57976	nuclear migration
24	177	2.0	1974	2 T16703	hypothetical prote
25	176	2.0	2843	1 RBHUAP	adenomatous polypo
26	176	2.0	2954	2 T14156	kinesin-related pr
27	175.5	2.0	1677	2 T43021	vitellogenin precu
28	175.5	2.0	2269	2 T28677	roptory protein -
29	174.5	2.0	3092	2 S46009	GTPase-activating

30	173.5	2.0	1744	2 F86161	F1003.10 protein -
31	173	2.0	1063	2 D86731	hypothetical prote
32	172.5	2.0	2232	2 T34434	hypothetical prote
33	172	2.0	895	2 T22061	hypothetical prote
34	172	2.0	2845	2 T49505	adenomatous polypo
35	171	1.9	2819	2 A90551	conserved hypothe
36	170.5	1.9	1957	2 T38077	hypothetical coile
37	170.5	1.9	2663	1 S28261	centromere protein
38	169.5	1.9	1238	2 T40120	C2H2 type zinc fin
39	169.5	1.9	4128	2 JC6306	protein kinase [EC
40	169	1.9	897	2 A54696	EGF receptor subat
41	169	1.9	1837	2 T41023	probable nuclear p
42	169	1.9	1858	2 T18273	i-phosphatidylinos
43	169	1.9	1916	2 S46157	RIF1 protein - yea
44	168.5	1.9	2253	2 T30336	nuclear/mitotic ap
45	168.5	1.9	3225	2 I52300	giantin - human

ALIGNMENTS

RESULT 1

JC5984

CAMP-dependent cell aggregation protein - slime mold (Dictyostelium discoideum)

N:Alternate names: AmiA protein

C:Species: Dictyostelium discoideum

C>Date: 06-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004

C:Accession: JC5984

R:Nagasaki, A.; Sutoh, K.; Adachi, H.; Sutoh, K.

Biochem. Biophys. Res. Commun. 244, 505-513, 1998

A:Title: A novel Dictyostelium discoideum gene required for CAMP-dependent cell aggregat

A:Reference number: JC5984; MUID:98189212; PMID:9514897

A:Accession: JC5984

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-1148 <NAG>

A:Cross-references: UNIPROT:O77203

F:316-338/Domain: leucine zipper #status predicted <LEU>

F:391-401/Domain: transmembrane #status predicted <TMM>

Query Match 13.9%; Score 1225; DB 2; Length 1148;
Best Local Similarity 30.3%; Pred. No. 6.5e-61;
Matches 321; Conservative 228; Mismatches 379; Indels 132; Gaps 28;

Qy	56	KLGHNNFTKLCIDIGHSEBKIGHYEDIIICRLALLNEAKVRAAGRALRYLIODSS	115
Db	68	KLQHLNKLVEKKSI-PDLNKLGISPTQLYKSVRPPFALPPKTIITAGLRVMTYLSNSN	126
Qy	116	ILQKVLKLVYLIARCIDIQSNEVERTQALRLVRKMITVNASLFPSSVTNSLIAGND	175
Db	127	NVKELDLKVQYFITSRLDRDKHSEPERIQALKIIRTIMEIDCSLMPHCFVKGLVSAEN	186
Qy	176	GLQERDMWRACATICEALONPEVALRGGLNTILKNVIDCOLSINEALITILHL	235
Db	187	---QEDNFCVCLCELTESIRNPQISSHCGGINTVFDAVLDPFYQIQESLLCILYLL	243
Qy	236	NHPKTRYVRADEVLELILAPYTD-PHY-----RHSPDTAEGOLKEDEA--RFLASMG	287
Db	244	SDKDTRIYIRPKSDLEIILAPLNSFNIGVKLCKASKEKEKEKEDEAVAMKKATASKA	303
Qy	288	IIATFRSWAGIINLCKPGNSGIQSLIGVLCPINMEIRRGLEVLVDYIFRLPLPVVTEFI	347
Db	304	VLTLIKSWIGIISL-NSDDQGLKSVDTLRMPQIELQEKALDSIFEIFRVLPKSIQE--	360
Qy	348	EALLSVDPGRFQDSWRSLDSGFVAEAKTILPHRARS-RPDLMDNYLALILISAFTRNGLLE	406
Db	361	----TFGPKATQTFNF-----GSETLQDLPSRTRSLRHNLNNYLSVLLIAFDINGLIE	411
Qy	407	GLV---EVITNSP-----DHTSVRATILGELLHWANTILPHSHSHHLCPLTM	453
Db	412	GLVTLGNYVANRDMGMSQEKCKSKNLSKSTVLLAEILLHMSNALLPPSQCAKQLTSLV	471
Qy	454	NMAASFDPKEKRLRASAALNCLKRF-HEM-----KKRGPKPYSLHL	494

Db 472 NSAGSFRDPTLTSSNTMTVNLHSHYSHNLSSSTLMDSTLAIGLTGANKWRRIQDQRRRL 531
Qy 495 DHI--IQKAIATHQKQDQVLRVQKQDIFILKDTFEEALLNLRDSQVLOKHENLEWNNLIG 552
Db 532 DKVDDVKMKQWHEMDNQFQO-----KIKDTQVLTVDKYQKSWELMF 574
Qy 553 TILKPNVNLVRYKQDQLHRFVRRRLIYFYKPSKLYANLDDFAKAKOLTVVGQOFTFL 612
Db 575 ELLEGFLNPPQHLNLTKTFKIKILSLFLRPKNKLFSTMAWTENLKYVR-TACVALEVL 633
Qy 613 LSEBEGQVLELDVKDIVQ-----WLNASSGMKPE--RSLQNNGLLT 653
Db 634 I-SHEIGFDFLKD-NKTIITQIDMLKVELDYNIKPPSSSSSSSKKONVRLNPEKVLK 691
Qy 654 TLSQHYFLFTGLTSLCHPHGVKMLEKSVFQCLNLCSLKQDHLKLTVSSLDYGRDGLA 713
Db 692 TMSREYFTWGLTSSNLGLLEILARNNIYDYIKPLAELPGRDDLHLILTSLDYNVNGAS 751
Qy 714 RVLSKILTAATDACRLYATKHLRVLRLANRVEFFNNWGLIELLVQTLHDKNKTISSEALDI 773
Db 752 RTILQKLTSSSRVRYLATKYLRFLRSVGQDFSNWGVLLVQQLNDVDKVSALSINV 811
Qy 774 LDACEDKANLHALIQMKALSHLGDGLLLLRFLSLPKGF-SYLNERYGVAKOLEKWH 832
Db 812 LDACDDPSCLEVLDLKNLKLKPKGSKSLLRFLSSPAGLENLLQNGFVQEQLWI 871
Qy 833 REYNKVDLIERQLEALT---TYRKPFVGDGNYVRRSNQRLQRPVHYLPILHYGOLVH 888
Db 872 TSENATVNAIESAVSESLSPSVWRFKEAPDGSS-----TSGVYLPHPFFGELAK 921
Qy 889 HKTGCHLLEVNQIITELCRNVRTP---DLDKWEEIKKLKSLWALGNIGSSNWNLLQ 945
Db 922 TERGCQLIRSNRYORFLKIOTPAKQLDK-----RASLTAIGHIGSSVDGYSFVKE 974
Qy 946 ENVIPDILKAKQCEVLSIRGTVCVYVGLIATKQKQCDILKCNWDAVHRSHKHLMPVVP 1005
Db 975 SDTIKLIGHIAKQCALASTCFYALGMISCIIEBAQPIFNSFGWSPSDLSNRI--LLP 1032
Qy 1006 DDVEQLCNEISSPTLSLNSSTSSRHNSSESVPSSMP 1045
Db 1033 KDLKN--STLSVP-----QVQYQGSWADHSFETLPSNH 1065

RESULT 2
T39379
Sexual differentiation and meiosis protein ste20 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39379; T46557
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Raper, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, September 1995
A:Reference number: 221849
A:Accession: T39379
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1309 <DEV>
A:Cross-references: UNIPROT:Q09743; EMBL:Z54140; PIDN:CAA90815.1; GSPDB:GN00066; SPDB:SH
A:Experimental source: strain 972h-; cosmid c12c2
R:Hilti, N.; Baumann, D.; Edenharter, E.; Stalder, M.; Schweingruber, M.E.
submitted to the EMBL Data Library, February 1998
A:Description: A new gene, ste20, in fission yeast Schizosaccharomyces pombe - necessary
and pH regulation.
A:Reference number: 223064
A:Accession: T46557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1309 <HIL>
A:Cross-references: EMBL:AJ223984; PIDN:CAA11758.1
C:Genetics: ste16, SPBCL2C.02c; ste20
A:Map position: 1
C:Function:

A:Description: necessary for sexual development and also responsible for sensitivity, met
Query Match 11.2%; Score 989.5; DB 2; Length 1309;
Best Local Similarity 27.1%; Pred. No. 1.6e-47;
Matches 314; Conservative 243; Mismatches 455; Indels 147; Gaps 39;

Qy 19 NDSEENVPLDLTREPSSD-----NLREILQNV--ARLQVSNMRKLGHLNANFTKLLC 68
Db 189 NSSSEDQPNG--KRFSLSQNSFSQPLNPLWLDNIYKACLEG-SMKDVIDSSNNLCYELH 245
Qy 69 DIGHESEKLGPHVEDIIICLRALLNEAKVRAAGLRALRYLIQDSSILQKVLKVDYL 128
Db 246 E--HSDPAYAKNFSLIPTTILSMLELVNVEVTSVYVRLLEHLFLDATALFSCOMLNPWI 303
Qy 129 IAR-CIDIQOSNEVERQALRLVRKMITVNASL-----FPSSVTNSLIATVNDGLQERDRM 183
Db 304 LSKSLSGTDAYQIEREQAFRLNTLYFLSTEGHEDYLSGITITLISI-CEHVSQVSRG 362
Qy 184 VRACIAIICELALQNPVVALRGGLANTILKNIDCOLS-RINELALITLHLHLNHPKTRQ 242
Db 363 I--AVEITLIELMIIRPKILFKANGLRVLMTSLIDGISENLAASAALVYLLDDPESAC 420
Qy 243 YVRADVELERILAPYTDHYRHSPTAEGQLKEDREARFLASKMGITATFRSWAGIINLC 302
Db 421 YVNLFDIGILLSPFTSSSRDFTFNSSESEQEAQAKWSAKVASV-LNNSWGLLALS 479
Qy 303 KPGNSGIQSLIGVLCIPNMEIRGLLEVLDFRLPLFVVTVEEFIEALLSVDPGRFODSW 362
Db 480 TNDQALRSIVDTLRVSPSPAPRSVIDLFFLIQFVE-----YSSW 519
Qy 363 RLSQGFVAEAKTILPHRARSPP-----LMDNYLALILSAFIRGLL 405
Db 520 --SESFLAGKRLTVVQNAQVNDNDNNINNVIPDGSNNKYMSSLRQHTAVLLFLEIGLV 577
Qy 406 ECLVEITNSDD-HTSVRATILGELLHMTILPHSHSHLHCLPTLMNMAASFDIPKE 464
Db 578 ESIVCMIRASDDPSASRKATYLLGEVLELSDELPIHLGAKIQLSPSLFNWASQF--TAE 635
Qy 465 KCLRASAALNCLKRFHEMK-KRGPKPYSLHLDHIIQKAIATHQKQDQVLRVQKQDIFILK 523
Db 636 DRFVATSVLQSLIESLRNVKPHSATQPFQSTTSLAFKE-----OKTDSFGROQVEHYK- 689
Qy 524 TEEALLIN-----LRDSQVLOKHENLEWNNLIGTILKPNVNLVRYKQDQLH--RFV 574
Db 690 LKMGQIDDSHFRSMLAETNVLATKYNQKRWDTLVUQIWEGLPSPKI-DETLRTKFM 748
Qy 575 RLLLYFYPSSKLYANLDDFAKAKOLTVVGQOFTFLSEEDQGYLED--LVKQIVQ 632
Db 749 RLLAFYKPFNSFSSIQ-NTKPNQKFIKVGCLVFTLLANPE-GVKYLSKVIKQIAE 806
Qy 633 WLNASSGMKPE-----LQNNGLLTLSOHYFLFTGLSCHPHGVKMLEKSVFQCLNLC 689
Db 807 SLSDIDGYSEQVSEPIFNSRLQKLTHTGYFPMPLVLSQKEGHAIMERWRITFTLYHLT 866
Qy 690 SLKNOHDHLKLTVSSLDYSRDCGLARVILSKILTAATDACRLYATKHLRVLRYANV--EFF 747
Db 867 ELNRDRLIITFLNLDVYLEGHTLIIFSKALNTQCAVRLTATKHLAALINSEANDNL 926
Qy 748 NNWGIELLYQLHDKNKTISSEALDILDEACEDKANLHA-LIQMPALSHLGDGLLIL 806
Db 927 NHWAILLIFQLVDPCLVEYCKTAVKLVNEVCARNENLLAQVVQLQPSLAHLGIEISPLLL 986
Qy 807 RFLSIPKGFSYLNERGVYAKOLEKWHREYNKVDLIERQLEALTYYRKPFVGDGNYVRR 866
Db 987 RFLATTGVGFYLSSEINFIEHLDNWNVHRNIDYVLDLEQNFPLSFVSNLKIIDKKNNEPD 1046
Qy 867 SNQRLQRPVYLPILHYGOLVHHKTKCHLLEVNQIITELCRNVRTPDLK---WEBIKKL 923
Db 1047 EN-----ILPLHYFGEILVKSPOCEVLESSGHF-ESFMGTLVFEYDKPLGNEAIRQL 1097
Qy 924 KASLWALGNIGSSNWNGLNLOEENVPIIDLKAKQCEVLSIRGTVCVYVGLIATKQKQCD 983
Db 1098 KSALWALGNIGTKDQGITFLINHDTIPLIVKVAENSLIPVRGTAYFVLGLISRTSGVE 1157

984 ILKCHNDVAV-----RHSRKHLPVVPDDVEQCNELSSIPSTLSLSESTS 1030
1158 ILESILWYSLMLGMSQICIPRHAGQVL--STPRNVFVNRVPTPEPSSLLSLT- 1214
1031 SRHNSSESVPSSMFILDDRFSGSSSTFFLDINEDETEFTFYDRSGPIKDKNSFPFAS 1090
1215 ---NSEREVIRL-----VSNLSNHLV-TNESAR-----QLTKIRSKNA-KVPSS 1253
1091 SKLVKRIILSLTLPNKXH 1109
1254 KRLVK-----ACMTILGKPH 1268

RESULT 3
S50596
hypothetical protein YER093c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: S50596; S59352
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of *S. cerevisiae* cosmids 9747, 8198, 9781, and lambda clones
A;Reference number: S50436
A;Accession: S50596
A;Molecule type: DNA
A;Residues: 1-1430 <DIE>
A;Cross-references: UNIPROT:P40061; EMBL:U18839; NID:G603313; PIDN:AAB64648.1; PID:G6033
R;Korch, C.; Mountain, H.A.; Wenzlau, J.M.
submitted to the EMBL Data Library, July 1995
A;Description: Structure and regulation of MET6, the vitamin B12-independent methionine
A;Reference number: S59352
A;Accession: S59352
A;Molecule type: DNA
A;Residues: 957-1430 <KOR>
A;Cross-references: EMBL:U32508; NID:G976437; PIDN:AAB60298.1; PID:G976438
C;Genetics:
A;Cross-references: SGD:S0000895
A;Map position: 5R

Query Match 8.28; Score 720.5; DB 2; Length 1430;
Best Local Similarity 23.18; Pred. No. 2.7e-32;
Matches 301; Conservative 218; Mismatches 531; Indels 255; Gaps 39;

18 RNDGSEEN-----VPLD---LTREPSDNLREILQNVAR----- 47
203 RNDGQNDSEELGDHSDLTQVLDDEYLTPTSGETRNSQNLNRNSTVNSRNNHS 262
48 -----LQGVSNM-----RKGLHAN-----FTKLLCDIG----- 71
263 TLSIPDLGSKNVNLTGDTTEKLDLENENQIPTSTTEATWLVSDYMQSFQEKVNP 322
72 -----HSEKLGHYEDIICRLALINEAKEVRAAGLRALRYLIQDSSI 116
323 FIAQANGVTLKEHSEIRKDLVLTFSMSIQNLNGLKTAASAYRVCYRLINSSIF 382
117 LQKVLKVDYLIARCDIQOSNEVERTQALRVKMITVNASLFPSSVTSNLI- AVGND 175
383 IDELELRLDPAFIISLAKNSQIEREQALKWRRIEYV-----NGVTQGMQAIIS 437
176 GLQERDMVRACIAICELALQNPVVALRGGLNTILKNVIDCOLSRINBALITLHL 235
438 VERPESLRLHMAETLELCFAPEMVMEKGRVIEGFLQDVTSFSLASVILDTILQLM 497
236 NHPKTRQYVADVELEIRLAPYTDPHYRHPDPAEGQKEDREARFLASKGIIATFRSW 295
498 ATHKTRQHFLEDFNVSVLLTVFSDTNTKSNVNERMQ-----NASTL-ISITLNSY 547
296 AGIINLCKPGNSGIQSLIGVLCIPNMEIRRGLEVLVDIIFRL-PLPVVTEFTEALLSVD 354
548 NGFMFNSNNFKPLQVLSFFQIP--ICAQYLDIFLDVLKIKPLPKYPR----- 595
355 PGRFQDSWRSLSDGFVAEAKTILPHRARSRLMDNLYLILSAFIRNGLLEGLVEVIT- 413

596 -GRSHSFK-----PIPSQYKCEMSVNQRLALIVILENSEFPVPHLELENE 642
414 -NSDDHLSVRATILGELLHMANITILPHSHSHLHCLPTLMNMAASPDIPKEKLRASAA 472
643 EDRDDHLVAKGRYLLTYFNLRMNLVDKKYT-----SVSPKPYKENTTYVNET 690
473 LNCILKRHEMKGRPKPYSLHLDHIIQKATATHQKRDQYLRVOKDIPILKDEBALLINL 532
691 FQFKIAYKNR-----NRNTIGSGIDYAGNIKSPSKNIKENTILLREVDFFRFRM 742
533 -RDSQVLQHKENLEWNNLIGTILKWPVNLRNYKD-EQLHRFVRRLLYFYKPSKLYAN 590
743 VYDSKVLQTKDTRWNWNIINELLEGLPLNKKQLEELVKSTKFRRLLVFVRLRSPFN 802
591 LDLDPAKAKQLTVVGCOFTPELSEBEDGQYLEDLVKDIVQ-----WLNASSGMMKPSRL 646
803 VNKAGAKUSQYVQGCQFFKTLTATPE-GMKILMDDTKIIPQASLMFRAMEGNISNIF 861
647 QNNGLLTTLTSLQHYFLFIGTILSCHPHGVKMLEKCSVFCQLNLLNCLKNQDHL--LKLTVSS 704
862 NKNKLEKIIFGYFKFIGILTQSKNGVHILTRWNFFTVYKMFQFESKLGLEFLLIPIE 921
705 LDYSRDLGLARVILSKILTAATACRLVATKHL-----RVLLRANVEFFNNGW 751
922 LDLKYSRCHRVIIIGKALVAVANEKVRIEATKHIGDKELLSKESDLKANKVKLQOFK 981
752 IELLVTLQDKNTTISEALDIDDEACEDKANHALQMPALSHLGDGKLL----- 803
982 MEMLTQLYD-----LSPSVAVADAQ-----LYECIVAGNGSEELTSFRMFLNQMVFI 1031
804 ---LLARFLSIPKGFYSYNERGVYAKLEKWHREYNSKYVDLIEQL--NEALTYRKPV 858
1032 RSPILFELLSRPYGFQLLNEINFVKEERDSWLSKKNLEYVHVEEFLKKNESI----- 1084
859 DGNVYRRSNQRQORPHVYLPILHYGOLVHHKTHGCHLE-----VONIITELCRNVRT 911
1085 ---NAKSLTFOQKSR---LPLHFYBSLTKTEDGILLSSQTGLVTPMNVKIVYVGNM 1137
912 PDLDKWEEKKAKSLWALNGNCSNWNGLNLOEUVIPDILKLAKECVLSIRGTCVYV 971
1138 ATVENAKEILDLAALWCVGFISGTGELGILLDNYSLVEDIIIEVAYNASVTSVRFTAFYV 1197
972 LGLIAKTQCCDILKCHNWD-AVRHSRKHLPVVPDDVEQCNELSSIPSTLSLSESTS 1030
1198 LGLISMTRECEILDEMGWCCVSVQDEPIGIALPNRLDRP-----LSVNEHKS 1247
1031 SRHNSSESVPSSMFILDDRFSGSSSTSTFFLDINEDETEFTFYDRSGPIKDKNSFPFAS 1090
1248 AFGEYSDEMI---VFNKSDG-----DLIEKCLPIEPDLDKLKEKDTAENPLN 1292
1091 SKLVKRIILSLTLPNKXHSRSDPKGKLSSESSTNRRIITLTPSV-DFNHSDDFTP 1149
1293 EKIITNKYNDIT-----SQTITVSGENSSLLFANEG-----LSSPYVTQYRNDSDSIE 1340
1150 ISTVQKTQLQETSFMGN---KHIEDTGST--PSIGENDLKFTKNF 1189
1341 SKVLHIVSQLGNHILSNHVAKEITEINNKYGPRLFENEKMFVKVF 1385

RESULT 4

T21539
hypothetical protein F29C12.2 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T21539; T27452
R;Steward, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19436
A;Accession: T21539
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-724 <WIL>

Query Match	4.1%;	Score 365;	DB 2;	Length 724;
Best Local Similarity	19.3%;	Pred. No. 1.1e-12;		
Matches	192;	Conservative 154;	Mismatches 346;	Indels 304; Gaps 33;
Qy	712	LARVLSKILTAATDACRLVATKHLRVLLRANVEFFNWGIELLVTLQHDKNKTTISSEAL	771	
Db	1	MVRVILGKALTSTSEARKWTRYLAVLASEFDLPTFSDWGIQYMLRQLADESSKVVPHI	60	
Qy	772	DILDEACED--KANLHALIQKPAHLGDKGLLLLLRLFLSIPKGFSYLWNGVYAKOLE	829	
Db	61	RILNRWLPPEPSRNLr-----KIDWSLFGAGDLLRAHVAFMES--ECASDEDEVDRDVR	113	
Qy	830	KWHREYNSKVVDLIEEQLNEALITYRPVGDNDVYRSNQRLQRPH--VYLPILHYQLV	887	
Db	114	FWMSDFNKYKLNIIIDEMKEMMFVKSIDGS--FSRSSDIPDTSIGVQAPLHLFAALG	171	
Qy	888	HHKGTGCHLLEVONIITELCRNVRTPDLDKWEEIKKLKASLWALGNIGSSNWGMLLQOEN	947	
Db	172	SHTGRIHLIDENVAEDLLSICKTKC--FEEL---KASLLAVSSIGSTDGGFEILPAD-	225	
Qy	948	VIPDILKAKQCEVLSIRGTCCVYVLGIATKQOCDILKCHNWDARVHRSKHLWPFVDD	1007	
Db	226	AVPVVIKIAEBHPVLTVRGIAFWALCTFSQCIEGAKRLGSGFWr-----	269	
Qy	1008	VEQLCNELSSIPSTLSLNSESTSRHNSSESVPSSFFILEDDDFGSSSTSTPFLDINED	1067	
Db	270	-----SHRFYAMDLSAKS--VVADDFGTPGGAT-----	297	
Qy	1068	TEPTFYDRSGPIKDKNSFPFPAASKLVKNRLINSLTLFNKHKRSSDPKGGKLSSESKTS	1127	
Db	298	-----PGSTCSSWRPARKITTIQSHRNSSMWDSQNNVKQSRAKE	336	
Qy	1128	NRRIRTLTEPSVDNFHSDDFPISTVQKTLQLETSPFGNKHIEDTGTPSIGENDLAFETK	1187	
Db	337	SALVRR-----GHSKG---RSRSOSTGDIHREVIAKRESRIDSPFNQRUNSEKLYLK	385	
Qy	1188	NFGTENHRENTSRRLVVESTSSHMKIRSQFNTDTTSGISMSWSSPSRETGVGDATT	1247	
Db	386	TSGTSDDSSSIPIYHGRVTNSSGHVQ-----DDTTAVTISPNHLNMEENV-----	432	
Qy	1248	MDTPCGSMSTVVSTKTIKTSHYLTPQSNHLSKNSVSLVPPGSSHTLPR-----R	1299	
Db	433	----CKS-----AATSRISTDRRANTTNSLPDEEE-----APKTRSSTVARCI	479	
Qy	1300	AQSLKAPSIATIKSLADCNFSYTSRDAFGVATLKLQOQ-----RMHPSLSHSEALAS	1353	
Db	480	TEQLEAEGVLA-DGIMEIHFSC-----RLREKHYLFAFRVACLQITRHVGD	525	
Qy	1354	PAKQVLFDTTITMKANSPESELTPSRFMKALSYASLDKLLSPINQNTLQSSSVSRSMV	1413	
Db	526	PVRVYFMTREBERHPADYRQVLRDPWL---YNELKKED-----NTVKKTINVVPIQ	574	
Qy	1414	SSATVGGSDDXIGIALPVDVI-----NDIIFQVKD-----IPYQTKNI	1464	
Db	575	T-----VALPTEIETMCNIIIPAKPKSDPIFSPHENDDASAGVDRGART-----	618	
Qy	1465	AGGIPSGTGGIVKNSFHILLROOMSITETWNSITHSDASLFLRSTEDTGHQEHETDONCLYCV	1524	

QY 381 ARSPDLMDNYLAI-ILSAFIRNGLLGLVE-----VITNSDDHISVRATILGELIHM 433
 Db 492 TERRKPTIDLLAARVLVSQFV---LINAQLPLSLARLILAMPDSSGLKATLLMADMLRQ 548
 QY 434 ANTILPHSHSHHCLPTLMMA-----ASPDIPKEK---LRASAALNCL 476
 Db 549 APSHVPGYRAPVLSMPTLVOSACESLSQSHAVAAINGTDFINTEQTFLLHATNAELVL 608
 QY 477 KRFE-----MKRGPKPYSLHLHHIIQKAIATHQ-KRDQYLR----- 513
 Db 609 NRFDKLAQSWIRASASTSTVLKESDLMLFSPHSEPIPLASTMKKSDDSSFRSGGSPFK 668
 QY 514 --VQXDF-----ILKD--TEALLI-----NLDSQVLQHENLE- 545
 Db 669 RNTSPDVYEESESGESETVLRGRYDKNALMLMDADDYDAAFNIRDT-LDDDEENAKK 727
 QY 546 -----MWNLLGTILKWPNNV-----LRNYKDEQLHRFVRRLLYFYKPSKLYAN 590
 Db 728 FFDVSTGAFDWNYYEYVE--NIEKGYEKLKFTNTIQVHISNMFQYISPEQEKHK 785
 QY 591 LDLPFAKQITVVGCQFTFELLESBEDGQYLEDLVKDI-VOWLNASSGMKPKERSLQN- 648
 Db 786 RCID----RQQVIVTKSLIKLLLRVFOK-EAHISDAYQPIFLKYVESFKALLEKRSNPNA 840
 QY 649 -----NGLLTLSQHYFLFIGTSLCHPHGVKMLEKCV 681
 Db 841 YPAARNLSYTNMLHFAIIGTFTWTSKGLSILTNAGV 877

RESULT 6

F90073
 hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
 C;Species: Staphylococcus aureus
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C;Accession: F90073
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yurawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001
 A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A;Reference number: A89758; MUID:21311952; PMID:11418146
 A;Accession: F90073
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2271 <KUR>
 A;Cross-references: UNIPROT:Q99QY4; GB:BA000018; PID:g13702612; PIDN:BA43752.1; GSPDB:G
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: SA2447

Query Match 2.5%; Score 217; DB 2; Length 2271;
 Best Local Similarity 22.5%; Pred. No. 0.0015;
 Matches 156; Conservative 95; Mismatches 272; Indels 170; Gaps 23;
 QY 1016 SSIPSTLSLSESTSSRHNSSESVSPSMFILEDDRFSGSSTSTFPFLDINEDTEPTFYD 1074
 Db 1222 STSESVSSTSESTSLSDSTSESGSTSTSS---LSNSTSGSASIST-STSISESTST- 1274
 QY 1075 RSGPIKDKNSFPFASAKLVKRNILSLTLPNKKHRSDDPKGKLSSEKTSNR----- 1129
 Db 1275 -----KSESVSTLSLMSSTSTSLNSLSTSLSDSTSDSKSLSTSMSTSDSISTSK 1327
 QY 1130 --RIRLTPEPSVDFNHSDDFT-----DFTPISTVQKTLQLETSPMGKNHIEDTGST 1175
 Db 1328 SDSISTSTSLSGSTSESESDSTSSRSKSDSTSMGISMOSTSGSTSTSTSTSLSDSTST 1387
 QY 1176 -----PSIGENDLKFTNKGTE--NHNRENTSRLRLVVE 1206
 Db 1388 SLSLSASMNQGVDSNSASQASNSTSTSTSESQSTSTSTSTSTSTSTSTSTSTSTSLSD 1447
 QY 1207 SSTSSHMKIRSQSFNTDTTTSIGSSMSGSPRRTVGVDAATMDTCGSMSTVTVSTKIKT 1266
 Db 1448 STSISKSTSGSTSTSTSLSGSESESDSQISITSAEESTS-ESASTSLSDSTSTSNSSG 1506

QY 1267 SHYLTPOSNHLSSKNSVSLVPPGSGHTLPRRAQSLKAPSIATIKSLADCNFVSYSRD 1326
 Db 1507 ASTSTSLSNSASASESDSS---TSLSDSTSSAQSSSESQSTASLSLSDSTSTSNR- 1562
 QY 1327 AFGVATUKRL-----QOORMHPSLSHSEBALASPAKDVLTFTITWKANS-----FE 1372
 Db 1563 ---MSTIASLSTSVSTSEGSTSESTSESTSTSLSDSQSTSRSTASGASSTSTSD 1619
 QY 1373 SRLTPSRFKALSYASLDKEDL-LSPINQNTLQBSVSRSMVSSATYGGSDDYIGLALPV 1431
 Db 1620 SRSTASSTSTMTSTSTSDSQSMSTSTSTSTSMSTSTSTSLSDSVSDST----- 1665
 QY 1432 DINDIFOVKDIPYQTKNIPPHDDRGARAFAHADAGLPSTGGILVKNFSHL-----LRQOMS 1498
 Db 1666 -----SDSTASSTSGMSVSLSDSTSTSTST 1692
 QY 1489 LTEMNSIHSDASLPLESTEDT-GLQHTDNDCLYCVCIELGFPQPNQJLSAICSHSDQF 1547
 Db 1693 ASEVMSASISDSQSMSESVNDSSESVSESNSES-----DSKSMGSGTSTVSDSG 1739
 QY 1548 DIPYSDWCEOTIHNPLEW--PSKFGISGCDGVSQEGSASSTKSTELLGVKTIPTDDT 1605
 Db 1740 SLSVS-----TSLRKSESVSESSLSGSGMSQMSDVSTSDSSLSVSTSLASSESVSESDS 1794
 QY 1606 PMCRILLRKEVLRVLVINLSSSVTKCHETGLLT 1638
 Db 1795 -----LSDSKSTSGTSTSTSGSL 1814

RESULT 7

T39903
 serine-rich protein - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T39903
 R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M. submitted to the EMBL Data Library, November 1998
 A;Reference number: Z21889
 A;Accession: T39903
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-534 <LYN>
 A;Cross-references: UNIPROT:O94317; EMBL:AL033534; PIDN:CAA22127.1; GSPDB:GN00067; SPDB:
 A;Experimental source: strain 972h-; cosmid c215
 C;Genetics:
 A;Gene: SPDB:SPBC215.13
 A;Map position: 2
 C;Superfamily: pig submaxillary mucin

Query Match 2.4%; Score 210.5; DB 2; Length 534;
 Best Local Similarity 22.6%; Pred. No. 0.00034;
 Matches 104; Conservative 79; Mismatches 182; Indels 95; Gaps 16;
 QY 1016 SSIPSTLSLSESTSSRHNSSESVSPSMFILEDDRFSGSSTSTFPFLDINEDTEPTFYD 1075
 Db 90 SATPSETNSVSPVSVDPATSQLPSS-----TSFPTSEVTPSSTE- 134
 QY 1076 SGPIKDKNSFPFASAKLVKRNILSLTLP-----NKKHRSDDPKGKLSSEKTS 1126
 Db 135 -----SSLLDPPSVSAAILPSPSTSVESISSSSLSSSDP----LTSSTFS 176
 QY 1127 SNRRIRLTPEPSVDFNHSDDFT---PISTVQKTLQLETSPMGKNHIEDTGSTPSIGENDL 1183
 Db 177 SLSSTSTSSQPSVSVST 229
 QY 1184 KFTKNGTGNHRENTSRLRLVVE--STSSHMKIRSQSFNTDTTTSIGSSMSGSP--SRET 1240
 Db 230 LRSLSLSTSGISPTSSSSSTSSLSLSSSSSTTASSSSSSSIISSSSSSSSFTSTST 289
 QY 1241 VGVDAATMDTCGSMSTVTVSTKIKTSHYLTPOSNHLSSKNSVSLVPPGSGHTLPRRA 1300
 Db 290 I8SSSSSSSPTSSS 347

Db 707 SSNCKLMTTFEQEIEELLHNTNIGTYKQV-----TTL-----EERTKNYE 748
 Qy 245 RADVELERILAPYTD-----PHYRSPDTAEGQ--LKEDREARFLASKMGI-IATF-----R 293
 Db 749 KTIHKHEQTVHLLKDNEMAAHRKHAADAQAQLRQENRILRDTSSRLQIEKEITYHREQQ 808
 Qy 294 SWAGIINLCKPGNSGIGSLIGVLCIPNMEIRRGLEVLVDYIPRLPLPVVTEEFTEALLSV 353
 Db 809 SOSILLNSLEPIKTNLER-----SEMEGRQRLERQLDDTVR-ELAAQRHRFQEE-----856
 Qy 354 DPGRFQDSWRSLSDGFAAAEAKTILPHRARSPDLMNDVLLALILSAFIR-----NG 403
 Db 857 -----BEKFRSINEFKQOETAIKLDEEK-QLADKWOAELTS--VREELAEKKNKNE 908
 Qy 404 LLEGLEVEITN--SDDHISV-----RATILGELL-HMANTILPHSHSHLH 447
 Db 909 LSKKLQELVLTPTLNDNHIITAANKRAREFELKLDQATVIESITRELAKT-----REHGQFY 965
 Qy 448 CLPTLMNMAASFDIPK-----EKLIRASAAALNCLKRHEMKR-GPKPYS 491
 Db 966 ----KMSQSAESEIKRLHELHGLVSKQEEIEKKLRSEA-----ELKTRISDLEAE 1013
 Qy 492 LHLDHITOKATATOKRDYOYLRVOKDIFILKDTREALINLRDSQV---LOHKENLENNW 548
 Db 1014 AMLSNVTEQSKTVNQS--COLKSAQDDL-----KSLLEKTEANCTIRLRSNTSLVE 1065
 Qy 549 NLIGTILKWPNNVLRNRYKDEQHRFVRRLLYFYKPPSSKI-----YANLDLDF 595
 Db 1066 SLNAAEVKYGAMIQHSAD--IQELTRYKAFFKANDDELNOLKSGRSLQAYDELRLSN 1123
 Qy 596 AKAKQLTVVGQCTFEFLLESEDDGGYLEDI-----VKDIOVOMLN-----ASSGMKPE 643
 Db 1124 AEAQKL-----LDKERESEKRVSDLHALNSNLHDQIEALASKLAVLASQSNPN 1173
 Qy 644 RSLQNNGLLTLSQHYFLFTGLSCHPHGVKMLEKSCVFQCLNLCSLKNQDHLKLTVS 703
 Db 1174 SSLNESAM-----DGDQSLNASGLTAAEEG-----RNEQLLKLI-- 1208
 Qy 704 SLDYSRDLARVILSKILTATADACRLVATYKHLVLLRANVPFNNGI--ELLVTQLHD- 761
 Db 1209 -----KFLRKEKD--LFAAK-LDLIAKAEARLISEHAIQOKKVDLNGY 1249
 Qy 762 KNTVTSSEALDILDEACEDKANLHALIOMK--PALSHGLDKGLGLLLPLSIPKFGSYLNE 820
 Db 1250 LNQERAKSOTDW-----SANKHEVLRKTIETLNAITDSNRILREERNALTLRVAELTD 1303
 Qy 821 R-GYVAKQ--LEKWHREYNSKYVDLIEBQALNEALTYRKPVDGDNVRRSNQRLQPHV 876
 Db 1304 RISSVEKELPFLQCSNKELTSKI-----BEINVENTSLR--TEAIKWRQANA-----1349
 Qy 877 YLPHTLYGQLVHHKGTGCHLLEVOQNIITELCRNVRTPLDKWEIEIKKLKASLWALGNIGSS 936
 Db 1350 -----LVEKSN-----RNP-----BEFKRLOAREHLAKULLTA 1377
 Qy 937 NWGLNLLQENVI-----PDILKLAQCEVLSIRGTCVYVLGLIAIKTQGCIDI--LKC 987
 Db 1378 EKELNKKQSDBLTVLKQRMWTEIPMLNKQMIILD-----EARKQVDEFTNLKQ 1426
 Qy 988 HNWDAVHRSKHLWPVVDVQELCNELSSIPSTLSLNSSTSSRHSSESESPVSMFIL 1047
 Db 1427 NN---TRQQT-----DIMELKNRL-----LQKEEELKANEELT-----1458
 Qy 1048 EDDRFGSSSTPFLDINEDTEPTFYDRSGPIKDKNSFPFPFASKLVKNRLNSLTLPNK 1107
 Db 1459 -----KDKT-----IADKETKELQRLAKRYK 1481
 Qy 1108 KHRSSDPKGG-----KLSEKSTSNRRIRTL-----TEPSVDF 1141
 Db 1482 DFYIGLQSGGGTESAAELEKVRSELEVVNNQLRALKDEHEKITKECDEVKKTEPETDT 1541
 Qy 1142 N-----HSDFTPTISTVQKT--LOLETSFMGNKHIEDTGSTPSPGENDLKFTKNFGT 1191
 Db 1542 SAIRQBYKAKLDKLVLDLTVARTDLVNQETTFAGTKSSYD--ETIARLEKELQ-----1592

RESULT 10

JC5837
 364K Golgi complex-associated protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
 C:Accession: JC5837
 R;Tokii, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.
 Cell Struct. Funct. 22, 565-577, 1997
 A;Title: Identification and characterization of rat 364-kDa Golgi-associated protein re-
 A;Reference number: JC5837; MUID:98093490; PMID:9431462
 A;Accession: JC5837
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-3187 <TOK>
 A;Cross-references: UNIPROT:Q63714; DDBJ:D25543; NID:G516825; PIDN:BA05026.1; PID:G5168
 C;Comment: This protein plays a role in the formation and maintenance of the characteris-
 F;49-549,524-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predict
 F;3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 2.2%; Score 197.5; DB 2; Length 3187;
 Best Local Similarity 19.9%; Pred. No. 0.031;
 Matches 372; Conservative 296; Mismatches 736; Indels 469; Gaps 90;

Qy 2 AATGRGSLKRLVGRNDSGEENVPLDLTRPESDNLREILQNVARLQGVSNMRK---LG 58
 Db 1425 ALISRKALK-----ENKSLQEQ--LSSARDAVEHLTKSLADVESQVSNQEKDALLG 1476
 Qy 59 HL-----NNFTKLCDIGHS---EEKLGFHYEDIIICLRAL--LNEAKEVRAAGLRALR- 108
 Db 1477 KLALQERDKLIVEMDKSLLENQSLGSCES---LKLALGLLTEDKELMKLESVRC 1532
 Qy 109 -YLIOSSILQKVLKLVYLIARICIDIQOSNEVERTQ-----ALRLVRKMTVNASLFP 163
 Db 1533 SKTAESTEWQKKELQKEVEVLQSVENVNAERLQHVVSVRQKQEVYAKLRSAES 1592
 Qy 164 SVTNSLIAGVNDGLQBRDRM-----VRACIALICEALQNPVEVALR---GGINTILK 213
 Db 1593 DKREREKQL-QDAEQEEMKEMKRWKFAKSKQKILEEENDRLRAEAQPVGGANESME 1651

QY 214 NVIDC-----QLSRIN-----EALJ-----TTILHL-----LNHPKTRQ 242
Db 1652 ALLSSNASLKEELERITLEYKTSKEFEALMAEKNTLSEETRNKLQVAEQELKQASLET 1711
QY 243 YVRADVELERIIAPYTDTHYRSPD---TAEQQLKEDREARFLA---SRMGIIATFRSWAG 297
Db 1712 TEKSD-EPKDVIEEYEAUVGKSQSDSLSENAKLEDAEATLLANSAPGVSETSSHDD 1770
QY 298 IINLCKPGNSGOSLIGVLCIPNMEIR--RGLEVLVYDIFRPLPLPVVTEEFTE-----348
Db 1771 INNYLQQ-----LDQLKGRIAELEMEKQKDRSLSQTLSEKNAKLLAQISAKDSELKLLBEE 1826
QY 349 -ALLSVDPGRFODS-----WRLSDGFVAABAKTILPHRBSR-----PDLMDNVLAILISAF 399
Db 1827 VAKINMLNQIOBELSRVTKLKE--TAEEKKDOLBERLMNQLAELNGSIGNYQDVTDQAQ 1884
QY 400 IRNGLLEGVLVEVITNSDDHISVRATITLIGELHFMANTIL-----PHS--HSH 444
Db 1885 IKNEQLESEMQLKRCVSELEBEKQQLVKEKTVSEIRKEYMEKIQAQKGPQSKIHAH 1944
QY 445 HUHCLPTLMMAASFDIPEKKL-----RASAAINLCRPHHEMKKGPQPY 490
Db 1945 ELQ-----ELLKEQKQEVKQLOKQOCIRYLGRISALEKTVKALEFVHTESQKDL 1992
QY 491 SLHLDHIIQKATATHQKRDQYLRVOKDIFILKDTTEALLINLRDSOVLOKHENLEWNNL 550
Db 1993 DATKGNLAQ-AVEHHKKAQELSSFK--ILLDDTQSEAAVLADN--LKKKKELQSNKES 2047
QY 551 IGTILKWPNNVL--RNYKDEQLHRFVRRLLYFPKPSKLYANLDDLDFAKAKOLTVVGCQF 608
Db 2048 IKSQIKQKDEDLLRLLEQAEKHR-----KEKNQKQKDALHREKAHVEDTLARI 2098
QY 609 TFFLESEEDG--QYLEDLVKDIVQWLNASGMPKERSLQNGLLTTLSSHLYFLPTGT 665
Db 2099 QVSLTRKDKMKELQSLDSTLAQAAFTKSSSLQDDR-----DRVIDEAKWEQRFQDA 2154
QY 666 LSCHPHGVKWLKCSVFCMLLCSLKNQ-----DHL--LKLTVSSLDYRDLGLARVILS 718
Db 2155 IQTKEEVEELKEE-----NCTALQKQRLQWTHMEELKITVSRLEHDKS-----IWES 2202
QY 719 KLITRAATDARLY-----ATKHLRVLRLANRVFFNNWGIELVLTQLHKNKNTISSEALDI 773
Db 2203 KAQTELHQHQKAYDKLQENKELMSQLEBAGQLYHDSKNEL--TKLESELKLDQSDTL 2260
QY 774 ---LDPAEDCKANLHALIQMKA-----LSHLGDKGLILLLLRF 808
Db 2261 KNSLEKREHENNLEGIIRKQEAQIQNCKFNCQLETDLTASRELTTRLHDBRINVKQEKI 2320
QY 809 LSIPKG-----PSYLAER-GYVAKOLEKWHREYNSKYVDLIE-----QLNEA 850
Db 2321 ISLLSKERAIQVAIAELHQHQSKEIKELNLLSQEENLLEENKKAVEKTKQLTAA 2380
QY 851 LTTYRK-----PVDGDNVYR-RSNQRLQRPVHYLPPIHLYQLVHHKGTGCHLLEVNII--902
Db 2381 LETIKESLEQAQLDSFVKSSSLQDDRDI---VSDYRQLEER---HL-----SVILE 2429
QY 903 -TELCRNVTPDLDKWEERIKKLKASLWALNGTSSWGLN--LLQENVIPDI--LKLAK 957
Db 2430 KDELQIDAAAENNLKKEERIGLRGH---MDDLSENKALDAELIQYRRDLNEVITIKDSQ 2486
QY 958 QCVEL-----SIRGTCVYVGLLIATKQGGCDILKCHNWDVAVRSHKHLWFPVPPDVE 1009
Db 2487 QRLLEAQIQONKELNBEVCVLEGRKGS-----EAEKQSLQMSLDALQEEHQ 2534
QY 1010 QLCNELSSIPSTI-SINSESTSRHNS-----SESVPSMFILEDDRFGSSSTSTFFLDI 1064
Db 2535 GLSKEIKSFEBQLTALHEGALAVYHAQRLVREEEVOKLTAAL-----SSSQKRTVDL 2587
QY 1065 NEDTETFFDRSGPIKDKNSFFPFASKLKVNRLNSLTLPNKKHSSSDPKGKLSSES 1124
Db 2588 QEELVCVQKEASKVSE-----IEDKLKREL-----KHLHNN-----AGIMRNET 2627

QY 1125 KTSNRIRITLTPSPVDFNHSDDFTPISTVOK-----TLQLETSFMGNKHIEDTGTSPGIG 1179
Db 2628 ETAERVAELARDLVEBEQK-----LLTVTKENKDLTAQIOAFGKSMSSLOQSRDHATEE 2682
QY 1180 ENDLAKFTQNGTENHRENTSRERLWSESSTSHMKIRSQSFNTDTTTSIGISMSSPSRE 1239
Db 2683 LSDLKKKYDASLKELAQKGRQDLGRES-----VLSQAAPPLTTSNISRLKLNQO 2736
QY 1240 TVGVDAITI--NDTOGSMSTVVSTKTIKTSHYLTPOSNHL--SLSKSNVSVLVPPGSSHT 1295
Db 2737 LISKDEQLHLSSLESHSNVQVSFT-KAMTSLQNERDHLWNELEKFRK-----SEEG 2788
QY 1296 LPRAQSLKAPSIAITIKSLADCNFSYTSRDAFGYATILKRLQOQ-----1339
Db 2789 KQSAAPSAASPAEVQSLKKAMSSQLQNDRL-LKELKNLQOQVLMNQHEITELRPLKA 2847
QY 1340 RMHPSLSHSEALAPAKDVL-----FTDTYTKANSFE--SRLTSPRFMKALSYASL 1389
Db 2848 QLOESQDQTKALQMKELRQENLSWOHELDQLRVEKNSEWELHERRMEKEQVLMALS-----2903
QY 1390 DKEDLLSPINQNTLQ--RSSSVRSVMV-----SSATVGGSD--YIGLALPVDIN 1434
Db 2904 DRQOQLGHL-QNLURELSSSQTLPTQYORQASSGTSALDGSQNLVYETDCLRTQLN 2962
QY 1435 DIFQVKDI-----PYFQTKNIPPHDDRGARAFADH-----1464
Db 2963 D--SLKEITHQKELRIQQLNSKFSQLLEKNILSTQLRDANQSLRDSQHHYNSLNFHNCAL 3020
QY 1465 -----AGGLP-----SOTGLVKNSFHLRQOQSLTMSI 1496
Db 3021 ERQVOQLQATGPLNADVAPGAQPKQKGIHIKTETTTGEEQPSFSEVQOQ-----LCNTK 3075
QY 1497 H--SPASLFLSTEDTGLQHTDDNCLVCVCEILGFQPSNQLSA-----ICSHSD--1545
Db 3076 HDLSKKLLBEED--QRLTAENALSLAKEQIRREHSEWESARTPIIAGCSQEQVL 3132
QY 1546 FQDIPYSWCQET 1558
Db 3133 LMDLPSS-CRRT 3144

RESULT 11
T34513
hypothetical protein ZK783.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34513
R:Ravell, A.; Vaudin, M.
submitted to the EMBL Data Library, August 1994
A:Description: The sequence of C. elegans cosmid ZK783.
A:Reference number: Z21536
A:Accession: T34513
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3507 <FAV>
A:Cross-references: UNIPROT:Q23587; EMBL:U13646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK.
A:Experimental source: strain Bristol N2; clone ZK783
C:Genetics:
A:Gene: CESP:ZK783.1
A:Map position: 3
A:Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 605/1; 1175/3; 1207/1; 1409/2; 1504/1
3504/1

Query Match 2.2%; Score 197; DB 2; Length 3507;
Best Local Similarity 19.3%; Pred No. 0.039;
Matches 186; Conservative 138; Mismatches 335; Indels 304; Gaps 42;
QY 902 ITELCRNVTPDLDKWEERIKKLKASLWALNGTSSWGLNLOENVIPDILKLAQCEV 961
Db 1802 IEERPESDKTACTDWSRLCELEKQCTV-----DDEEVP-----QCGA 1840
QY 962 L-----SIRGTC--VYVGLIATKQGC-----DILKCHNWDVAVRSHKHLWP-----1002

Db 1841 CLPGHPINGTCQSIQISGLCAQ-KNDCNKAECIDI-----HPDSHFSCPDGF 1889
QY 1003 -----VVPDDVQLCN-----ELSSIPSTL-SLN----- 1025
Db 1890 IGDGMICDDVDE-CNNAGMCDENTKCENTIGSFNCVCKGVDEKCVVDEKKQPNRE 1948
QY 1026 -----SESTSSRHNSSESVPMFLEDDRGSSSTFFLDINEDTPTFYDRSGPIKD 1081
Db 1949 KIEIDENSSNSGQEK-PTTKGIVSSTSATSSEST-----APHV-----TTSISS 1996
QY 1082 KNSFPFFASSLKVNRILNS-----LTLPNKKHRSDDPKGKLSSESKTSNRRI----- 1131
Db 1997 TTSKDMTSSKSPENVTHSSSEPVSSSSKSTTASEITVSTPSESASSAPLTSPPAT 2056
QY 1132 -----RLTPEPSVDFNHSDDFTP-----ISTVQKTLQLETSMGNKHIEDTGSTPS----- 1177
Db 2057 TTEVITESV-----KSTTPKESSSEITVKLSKSPVETSSVKSPSPSTTSQSPTS 2111
QY 1178 -----TGENDLKTKNFGTENHRENTSRRLVSES-----STSSHMKIRSQ 1218
Db 2112 TVPETSSTVLSSEAPVTSTPTVHTSSETKPSLSASSTTGDTNSTPTSTSLASVKST 2171
QY 1219 SFNTDTTTS-----GISMS-----SPPSRETV-GVDATTMDTDCGSMSTVSTKTKTSHYL 1270
Db 2172 SAPEGTASVAPVKLSLSPDVSPQSTKTFPATSESSIVQABETSSGTSVKSTSEPSHV- 2230
QY 1271 TPQSNHLSLKSNSVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFYSYTSRRDAFY 1330
Db 2231 -----TKLSITSNFSSVPVTPSKSTPVPESTQSTTPSG----- 2269
QY 1331 ATLKRQOQRMHPSLSHSEALASPAKDVLFDTITMKANFESRLTSPRM-----K 1382
Db 2270 -----QSLTPMNSNSELVTTSEPHVL-----SSSLSPDVQSSTTPNNLSSESSTVETPK 2318
QY 1383 ALSVASKED-----LISP-----INQNTLQSSSV-----RSMVSSAT----- 1417
Db 2319 TSEVLSNBEPTTEAPTTLSPDILSTTTNNLSQSSTVSTEDKSEISSENSEKPTSAPE 2378
QY 1418 -YGSDDYIGLALPVDINDIFQVKDIYFQTKNIPPHDDRGARAFADAGLPGTGLV 1476
Db 2379 LVTSVVTHVASSPDVTESEPDLLTGSSTENIP-----EASSKOTISSTTPDTTASEE 2435
QY 1477 KNSFHLLRQOMSLTEIMNSIHSASLPLESTEDGLQETDDNCLYCIEILGFQPSNQ 1536
Db 2436 PTRKSTMSPLSTT-----SNVLSSESTTPSSSKSPVSSSTEG----- 2474
QY 1537 LSAICSHSDFQDIPYSDWCQTHNPLE-----VPS-----KPSGIGCDGVSQEGAS 1587
Db 2475 ISVVTST-TEFSKVP-----ESTISSVLEEDLTKTPSPILEETTTASETSEPLTSLTV 2528
QY 1588 STKSTELLGVKTIPTDTPMCRILLRKEVLRLVNLSSSVSTKCHE--TGLLTIKEKYPQ 1645
Db 2529 SVRIHELTTSENVPKSES-----TTTSSSSKSPQBPAGILSTVVVPT 2574
QY 1646 TFDICLYSVSHLSHCHTFLPCRFQELFQDVQFLQMEHAEAVLATPPKQIPVDTS 1705
Db 2575 SSVSLITASEIAITSNTPK-----QGRTPITTSPLVKSTT 2613
QY 1706 AES 1708
Db 2614 SPS 2616

RESULT 12

S25370

MSB2 protein - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein G4017; protein YGR014w

C;Species: Saccharomyces cerevisiae

C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004

C;Accession: S25370; S64305

R;Bender, A.; Pringle, J.R.

Yeast 8, 315-323, 1992

A;Title: A Ser/Thr-rich multicopy suppressor of a cdc24 bud emergence defect.

A;Reference number: S25370; MUID:92383951; PMID:1514328

A;Accession: S25370

A;Molecule type: DNA

A;Residues: 1-1306 <BEN>

A;Cross-references: UNIPROT:P32334; GB:M77354; NID:g171993; PIDN:AAA34798.1; PID:g171994

A;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S64071

A;Accession: S64305

A;Molecule type: DNA

A;Residues: 1-1306 <RIE>

A;Cross-references: EMBL:Z72799; NID:g1322977; PID:g1322978; MIPS:YGR014w

A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:MGB2

A;Cross-references: SGD:S0003246; MIPS:YGR014w

A;Map position: 7R

C;Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase

C;Keywords: transmembrane protein

F;3-19/Domain: transmembrane #status predicted <TM1>

F;1189-1205/Domain: transmembrane #status predicted <TM2>

Query Match 2.28; Score 196.5; DB 2; Length 1306;

Best Local Similarity 20.94; Pred. No. 0.0087;

Matches 162; Conservative 119; Mismatches 270; Indels 225; Gaps 31;

QY 1004 VPDDVEQLCNELGISPTLSNGESTSSRH-NSESESVPMFLEDDRGSS---STST 1059

Db 157 VPID-----TSANSIPPTTSNIETTSAPLTSOTPLISTSTMSAADNVFSSAIPISASL 211

QY 1060 FFLDINDEBTFTFYDRSGPIKDKNSFPFFAFSSKLVKNRILN----- 1100

Db 212 TTTDSSSPQGT-STAGAIPOQSSADFPSSSEILVQSSADFPSPPTTTDLSAAPL 269

QY 1101 -----SUTLPNKKHRSDDPKG-----GKLSSESKTSNRRI-RLTE 1136

Db 270 QTSESSFTTASALP-----VSSDVGSSASPVVMSAAGIAQSSSTNDPTMSETFSL 325

QY 1137 PSVDFNHSDFTPISTVQKTLQLETSPMGNKHIEDTGSTPSIGENDLKFTKNFCTENHRE 1196

Db 326 TSTEVDGSDVSTVSALLSAPFLQTS-----TSNSFSIVSPSVFVPSQSSD--- 373

QY 1197 NTSRERLVBSSFTSSMHKIRSQFN---TDTTTSG-----ISSMSSSPSRRRTGVDA 1245

Db 374 -----VASSSTAN---VVSSSFSDIPPTQTSIGSVVVAQASALAFQSSTEVYGASA 423

QY 1246 TTMWDTCGSMSTVSTKTKSHYLPQSNHLSKNSVSLVPPGSSHT--LPRRAQSL 1303

Db 424 SS-----TMSLLSTTSLQST---TLDSSSLASSASSSSDLTDYGVSSSTASIPLLGASE 474

QY 1304 KA-----PSIATIKSLADCNFYSSTSS---RDAFGYATLKRLOQOQRMHPSL----- 1345

Db 475 QASTSSSFVSVPSFVPSQSSSDVASTAPSVSSFSYSTLQAGSGSMTPNPSSTIV 534

QY 1346 -----SHSALASPAKDVLFDTITMKANFESR-----LTPSRF 1380

Db 535 YSSSTGSSSEASASTASATLSSGSSSTMAGNLQSQPPSTSSLSSESQATSTAVLASSV 594

QY 1381 MKALSY-----ASLKEDLLSPINQNTLQSSS-----VRSMVSSATYGGSD----- 1422

Db 595 STTSPYTTAGGASTEASSLISSTSAETSOVSYSQSTTALQTSFASSTTTEGSETSSQGF 654

QY 1423 --DYIGLALPVDINDIFQVKDIYFQTKNIPPHDDRGARAFADAGLPGTGLVKNPSF 1480

Db 655 STSSVLVQMPSSISSEFS-----PSQTTTQMNSAS 684

QY 1481 HLRLQOMSLTEIMNSIHSASLPLESTEDGLQETDDNCLYCV---CIBILGFPQSNOL 1537

Db 685 SSSQYIISSTGILSQV-SDTSVSY-TTSSSVSQVSDTPVSYTTSSSSVSQVSDTPVSYT 742

QY 1538 SAICSHSDFQDIPYSDWCQTHNPLEVPSKFSIGSCDGVSGESASSTKSTELLG 1597

Db 743 TSSSSVSQVSDTPVSYTTSSSVSQVSDTPVSY---TTSSSVSQV-SDTSVPSSTSSRS 798

QY 1598 VKTIPDDTPMCRILLRKEVLRLVINL-----SSSVSTKCHE 1633
Db 799 VSOV-SDFVPTSSRSSVSQSSLSQPTTSSQRTTSTHGLSESSVSQASE 853
RESULT 13
T30531
A:agglutinin-like adhesin - yeast (Candida albicans)
C:Species: Candida albicans
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30531
R:Gaur, N.K.; Klotz, S.A.
Infec. Immun. 65, 5289-5294, 1997
A:Title: Expression, cloning, and characterization of a Candida albicans gene, ALA1, the
A:Reference number: Z20847; PMID:98053977; PMID:9393828
A:Accession: T30531
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1419 <GAU>
A:Cross-references: UNIPROT:O13368; EMBL:AF025429; NID:g2522219; PIDN:AA88
C:Genetics:
A:Gene: ALA1
C:Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
Query Match 2.2%; Score 192; DB 2; Length 1419;
Best Local Similarity 21.2%; Pred. No. 0.018;
Matches 164; Conservative 118; Mismatches 288; Indels 202; Gaps 33;
QY 1011 LCNELSSIPST-LSLNSESTSRHNGESVSPSSMPFILEDDRFGS--SSTSTFFLDIND 1067
Db 694 LSSDMPSSSTGLTSESSTVSYSDDSSSTIESSTLSSSDRCSISDTTSFWDSSSD 753
QY 1068 TPTTFYDRSGPIKDKNSFFPFPASSKLVRNRLNSLTLPNKKHRSSDDPKG-----GKLS 1122
Db 754 LESTSITWSSIDAQ-----SSHLVQS-VNSISSTQELSSSSSEESTPATDALVSS 805
QY 1123 E-----SKTSNRIRTLPEPSVDFNHS-----DDFTPISTVQTKTQLE-----T 1161
Db 806 DASSILSSDTSYYPSTSSDDPHTTAGESDLSISFITSVEISSDSVLSLSDPAS 865
QY 1162 SPMGNKHI-EDTGSTPSIGENDLKFKNFGT-----ENHRENTSRER 1202
Db 866 SFDSSSLNSDSSSPSQSDILTSSTFVLVPSFSLSSSSLSLTYPVNSSTYHA 925
QY 1203 LVVESTSHMKIRSGSFYDFTTISGISMSSPSPRETVGVDAITM-----DTCGSMST 1257
Db 926 SESESSVASPMASESANDDTYTLSESTDTTS-----SIGTDSSTVTCRRDNGDCIVT 981
QY 1258 VVSTKTI-----KTHYL---TPQSNHLGL-----GKSNVSVLVPPGSSHTLPR 1298
Db 982 GMPSSSIDSEQSDVTTTSFVASSTPTBAQSIITDNPIDISQTSASSTKSSVSVD 1041
QY 1299 RAQSLKPSIATIKS-----LADCNFSYTSRDA-----FGYATLRLQOQRMHPSL 1345
Db 1042 VVNSILLSETSLSSDDTSSTISSTTNSDGTGNINAGSSHTSTASIKESSIQKTGVTL 1101
QY 1346 SH---SEALASPAK---DVLFTDTITMKANSPEER-----LTPSPFKALS-----YA 1387
Db 1102 SSSYLSLKLSSDITIELITTELTIEDNEPNTFTTSPSSHSEIFSSDNLVSLK 1161
QY 1388 SLDKEDLL--SPINQNTLQSSSVRMSVSATYGGSDDDVIGLA-LPVDINDIFQ----- 1438
Db 1162 QVDRESTIKTPTDVTTVSSLSVHSTEASTATLGENSEFNASTPSNATLSLRSSTSS 1221
QY 1439 -----VKDIPYQTKNIIPPHDRGARAF-AHDAGLPGSTGGLVKNKSHLLRQOMS 1488
Db 1222 NHATESGTVKGEASAEAPSPPTSDNRLSYSTEAKGITVANSGSTNN-----L 1272
QY 1489 LTEIMNSIHSDALFLESTEDTGLQEHDTDNCLYCVCIELGFPQSNQLSAICSHSDFOD 1548
Db 1273 ITESVAAPATDSTSVL--IENPVVTTFDDNSSAAV-----DQPSKTKS----- 1314

QY 1549 IPYSDWCBCQTHINPLEV-----VPSK-----FSGISGSCDGVSEEG--S 1585
Db 1315 -----IEESIMNPDSTNETNNGFIATLSQAQVPSSTHSELISITTTAKTTDASMGDSA 1368
QY 1586 ASSTKSTELLGVKTIPTDDTPMCRILLRKEVLRLVINLSSSVSTKCHETGLL 1637
Db 1369 ASNSQPTTLIQVATSYNQ-----LITTVAGSSSATKHPSWLL 1408
RESULT 14
A41258
A:agglutinin core protein AGA1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: a-agglutinin attachment protein AGA1; protein N3431; protein YNR044w
C:Species: Saccharomyces cerevisiae
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
C:Accession: A41258; S63375; S17031
R:Roy, A.; Lu, C.F.; Marykwas, D.L.; Lipke, P.N.; Kurjan, J.
Mol. Cell. Biol. 11, 4196-4206, 1991
A:Title: The AGA1 product is involved in cell surface attachment of the Saccharomyces cer
A:Reference number: A41258; PMID:91304412; PMID:2072914
A:Accession: A41258
A:Molecule type: DNA
A:Residues: 1-725 <ROY>
A:Cross-references: UNIPROT:P32323; GB:M60590; NID:gl70963; PIDN:AAA34382.1; PID:gl70964
R:Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63346
A:Accession: S63375
A:Molecule type: DNA
A:Residues: 1-725 <POH>
A:Cross-references: EMBL:Z71659; NID:gl302551; PIDN:CAA96325.1; PID:e239834; PID:gl302552
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:AGA1
A:Cross-references: SGD:S0005327; MIPS:YNR044w
A:Map position: 14R
C:Keywords: glycoprotein; transmembrane protein
F:8-24/Domain: transmembrane #status predicted <TM1>
F:708-724/Domain: transmembrane #status predicted <TM2>
Query Match 2.1%; Score 188.5; DB 2; Length 725;
Best Local Similarity 22.7%; Pred. No. 0.0096;
Matches 107; Conservative 70; Mismatches 194; Indels 101; Gaps 15;
QY 1011 LCNELSSIPSTLSLNSESTSRHNGESVSPSSMPFILEDDRFGSSSTSTFFLDINDTTP 1070
Db 173 LSTTSSNPTTTLSTSTSPSTSTSPSTSTSPSTSTSPSTSTSPSTSTSPSTSTSP 228
QY 1071 TFDYRSGPIKDKNSFPFPASSKLVRNRLNSLTLPNKKHRSSDDPK---GGKLSSSEKTS 1127
Db 229 SSTSTSSSLTSSSTSTSPSTSTSPSTSTSPSTSTSPSTSTSPSTSTSPSTSTSP 288
QY 1128 NRRITLPEPSVDFNHSDDFTPISTVQTKLETSMGNKHIEDTGSTPSIGENDLKFTK 1187
Db 289 TSSYSTSTSPSLT---SSPTLASTSPSTSTSTSTSTSTSTSTSTSTSTSTSTST 337
QY 1188 NF-----GTENHRENTGRERLVVR---SSTSHMKIRSQSFNTD----- 1223
Db 338 SVSLSPSTPVYSPVSTSNVATPMTSTVETTVSSQSSEVITKSSISTIPFSMST 397
QY 1224 ---TTTSGISSM-----SSSPS-----RETGVGDAT-----TMD 1249
Db 398 YFTTVSGVTTMTVTCVPYSSSESTSLTSMHETVTTDATTCTVCTHESCMPSTSLITTSIK 457
QY 1250 TDCGSMSTVSTKTIKTSHYLTLPQSNHLSLSKNSVSLVPPGSSHTLPPRAQSLKAPSA 1309
Db 458 MSTKNVATSVSTVESSYACS-----TCAETSHSYSSVQTASSSVTQQTSTK---S 508
QY 1310 TTKSLADCNFSYTSRDAFGYATLRLQOQRMHPSLSHSEALASPAKDVLFDTITMKAN 1369
Db 509 WYSSMT-----TSDDFNKHATGK-----YHVTSSTGTSTSTSVSEATSTSSIDSE 556
QY 1370 SPESRLTPSRFMKALSYASLDKED-----LLSPINQNTLQRSSSVRSMVSSAT 1417

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2005, 15:10:51 ; Search time 75 Seconds
(without alignments)
7560.684 Million cell updates/sec

Title: US-10-782-244-3

Perfect score: 8811

Sequence: 1 MAAGRSLKSLNRVRGRND.....AEAVLPPKQPIVDTSAES 1708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281	3.2	56	14 US-10-029-386-32805	Sequence 32805, A
2	225	2.6	2297	14 US-10-245-802-20	Sequence 20, Appl
3	220.5	2.5	2344	9 US-09-815-242-12713	Sequence 12713, A
4	217	2.5	2271	15 US-10-282-122A-43924	Sequence 43924, A
5	214	2.4	2261	17 US-10-470-048B-60	Sequence 60, Appl
6	212	2.4	2319	16 US-10-408-765A-2075	Sequence 2075, Ap
7	211	2.4	2283	14 US-10-172-502-4	Sequence 4, Appli
8	204.5	2.3	1979	14 US-10-205-823-419	Sequence 419, App
9	204.5	2.3	2052	16 US-10-437-963-137285	Sequence 137285, A
10	200.5	2.3	1831	15 US-10-282-122A-71033	Sequence 71033, A
11	200.5	2.3	1870	17 US-10-470-048B-268	Sequence 268, App
12	197	2.2	3507	15 US-10-369-493-5784	Sequence 5784, Ap
13	195.5	2.2	1443	14 US-10-245-802-18	Sequence 18, Appl

14	193	2.2	560	9 US-09-815-242-13057	Sequence 13057, A
15	192	2.2	502	9 US-09-815-242-5904	Sequence 5904, Ap
16	191.5	2.2	469	15 US-10-282-122A-70453	Sequence 70453, A
17	190	2.2	454	15 US-10-424-599-232134	Sequence 232134, A
18	189	2.1	2224	14 US-10-115-563-14	Sequence 14, Appl
19	188.5	2.1	1725	15 US-10-369-493-2004	Sequence 2004, Ap
20	188	2.1	1270	14 US-10-437-963-159035	Sequence 159035, A
21	186.5	2.1	1724	16 US-10-437-963-159035	Sequence 159035, A
22	183.5	2.1	841	14 US-10-032-585-7122	Sequence 7122, Ap
23	182	2.1	2196	15 US-10-360-101-259	Sequence 259, App
24	182	2.1	2224	14 US-10-172-712-31	Sequence 31, Appl
25	182	2.1	2224	16 US-10-741-601-542	Sequence 542, App
26	182	2.1	2224	17 US-10-741-600-1561	Sequence 1561, Ap
27	180	2.0	2808	15 US-10-210-281-50	Sequence 50, Appl
28	179	2.0	22152	16 US-10-715-066-5	Sequence 5, Appli
29	178.5	2.0	5373	16 US-10-408-765A-741	Sequence 741, App
30	178	2.0	1881	14 US-10-032-585-7646	Sequence 7646, Ap
31	177	2.0	1974	15 US-10-369-493-5455	Sequence 5455, Ap
32	177	2.0	2621	16 US-10-437-963-122168	Sequence 122168, A
33	176	2.0	2843	9 US-09-987-482-1	Sequence 1, Appli
34	176	2.0	2843	15 US-10-392-113-21	Sequence 21, Appl
35	176	2.0	2843	16 US-10-408-765A-1970	Sequence 1970, Ap
36	174.5	2.0	3092	9 US-09-801-368-172	Sequence 172, App
37	174.5	2.0	3092	15 US-10-369-493-1470	Sequence 1470, Ap
38	174	2.0	2843	8 US-08-681-219-32	Sequence 32, Appl
39	174	2.0	2843	10 US-09-230-111C-30	Sequence 30, Appl
40	174	2.0	2843	14 US-10-092-138-30	Sequence 30, Appl
41	172.5	2.0	603	15 US-10-282-122A-72004	Sequence 72004, A
42	172	2.0	1384	16 US-10-473-576-22	Sequence 22, Appl
43	172	2.0	2036	16 US-10-437-963-170373	Sequence 170373, A
44	172	2.0	2568	9 US-09-866-108-3	Sequence 3, Appli
45	172	2.0	2568	16 US-10-723-361-13	Sequence 3, Appli

ALIGNMENTS

RESULT 1

- US-10-029-386-32805
; Sequence 32805, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: A60MICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32805
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008964.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
; OTHER INFORMATION: SWISSPROT HIT: P35659, EVALUATION 3.60e+00

Query Match 3.2%; Score 281; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.8e-12;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 196 LQNPVVALRGGLTILKKNVICQLSRINEALTTTILHLNHPKTRQYVRADVELE 251
Db 1 LQNPVVALRGGLTILKKNVICQLSRINEALTTTILHLNHPKTRQYVRADVELE 56

RESULT 2
US-10-245-802-20
; Sequence 20, Application US/10245802
; Publication No. US20030124134A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, John E.
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND METHODS TO VACCINATE AGAINST
; FILE REFERENCE: 013361.4003
; CURRENT APPLICATION NUMBER: US/10/245,802
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 09/715,876
; PRIOR FILING DATE: 2000-11-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent version 3.2
; SEQ ID NO 20
; LENGTH: 2297
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-245-802-20

Query Match 2.6%; Score 225; DB 14; Length 2297;
Best Local Similarity 21.1%; Pred. No. 6.2e-06;
Matches 157; Conservative 129; Mismatches 283; Indels 176; Gaps 29;
QY 1014 ELSS---IPST---LSLNGESTSSRHNSSESVPPSMFILEDDRFQSSSTSTFFLDINE- 1066
DB 874 ELSSRDIPSSIGLTSSESTVSSVDSYSSSTSESIASSVDSYSSSIESSTLSSDR 933
QY 1067 -----DTPTFFDRSGPIKOKNSPPFA-----SSKLKVRILNLSLTPNKKHRSSDP 1115
DB 934 YSSISDIT-TSFYDSSSDLESTITYSSSIDAQSSHLVQS-VNSISQTSQIBSSSEE 991
QY 1116 KG-----GKLSE-----SKTSNRIRLTLPSPVDFNHS-----DDFTPISTVQTLQ- 1159
DB 992 SSTSATLVSADASSILSSDTSYPPSTISPSDDFPHTIAGESDSQSIFITVTEIS 1051
QY 1160 -----ETSMGNKH-EDTGSTPSIGENDLKFTKNGFT-----ENHRENTSRERLV 1204
DB 1052 SDSVSLTSPSSFDSSRLNDSSTPDQDILTSSFFTLKSSGSRSSGTILS 1111
QY 1205 VSSSTSSHMKIRSQFNNTTGGISSMSSSPRFTVGVDATTMTDCCGSMSTVVSTKTI 1264
DB 1112 EESSDIPTTSTRY-SPSGMSSRRHYNTSTSTSVSDVVS-----SSVAGDETS 1159
QY 1265 KTSHYLTPOSNHLSKNSVSLVPPGSHLTPRAQSLKAPSLA--TIKSLADCNFSVT 1322
DB 1160 EESSVISESSSVTSVASESV-----ASESVASESVASESVAVSDISDLYT 1209
QY 1323 SSRDA-----FGYATLKELOQR-----MHPSLGSHSEALASPAXDVLFTDTITMKANSF 1371
DB 1210 TSEEVSTDSNMGSSPIPSQRSSIPIMSSDESSESSRESSTGILSEENSDIPTTF 1269
QY 1372 ESRLTSPRMKALSVASLKDPELLSPNQNTL--QRSSVRMWSVATYGGDDYI----- 1425
DB 1270 STRYSPSGMSSRRHYNTSTSTSVSDVSSVAGDETSESSVSVI SESSSVTSVASES 1329
QY 1426 -----GLALPVDINDIFQKIDIPYFQTKNI PHDDRGRARAHADAGLPCTGGVLV 1476
DB 1330 VASESVASESVAVSDISDLYTTSSEVVSTSDN-----SGMSSPI 1369
QY 1477 KNSPHLLRQOMGLTEIMNSIHSASLFLSEDTGLQHEHTDNCILCYCIEILGQPQSNQ 1536
DB 1370 PSS-----EQRSSIPWMS-SDESSRESSTGILSEENSDSI-----PTTF 1411
QY 1537 LSAICSHDFQDIPYDWCQEIHNPLEVPKPSG-----ISGCSDDGV 1581
DB 1412 STRYSPSGMSSRRHYNTSTSTSVS---DVSSSVAGDETSESSVSVI SESSSVTSVASEV 1468
QY 1582 QSGSASSTKSTELLGKVTIP-----DTPMCRILLRKEV-----LRLVINLSSV 1627
DB 1469 SESVASESVASESVAVSDISDLYTTSSEVVSTSDSKIVASTVSPSEQRSSIPIMSSSDE 1528

QY 1628 STKCHETGLLTI-----KEKYPQTF 1647
DB 1529 SSESRESSGTLSEENSDSIPTTF 1553
RESULT 3
US-09-815-242-12713
; Sequence 12713, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12713
; LENGTH: 2344
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12713

Query Match 2.5%; Score 220.5; DB 9; Length 2344;
Best Local Similarity 23.5%; Pred. No. 1.3e-05;
Matches 159; Conservative 93; Mismatches 292; Indels 133; Gaps 25;
QY 1006 DDVEQLCNELSSIPSTLSLNESTS-----SRHNSSESVSPSMFILF-----D 1049
DB 1167 DSTSDSTSTDSISEASIS-GSESTISLSESNSTSDSEKSAFLSESLSESTSE 1225
QY 1050 DRFGSSSTTFPLDINEDTEPT-----FYDRSGPIKDKNSFPFPASSKLKVRILNLSLTP 1105
DB 1226 SLSGSTSDSTLSLDSNBSGSGTSLNLSNSTSGSASISTSTSGSASTSVKSESV-STLS 1284
QY 1106 NKURSSDDPKGKLSSEKTSNPRIRLTLPSPVDFNHSDDFTPIST-VQTLQLETSPM 1164
DB 1285 TSTSTLSLSDSTLSLSTLSLSDSTSGSKNSL-----SASMSTSDS-----ISTRKSESLSASTSLS 1338
QY 1165 GNKHIEDTGSTPSIGENDLKFTKNGTENHRENTSRERLVVSESTSSHMKIRSQSFNTDT 1224
DB 1339 GSTSESESGTSS-----SESKSDSTSMLSMSQSTSGSTSVSTSLSDST 1385
QY 1225 TT-----SGISNMSSSPRETVCVGDATTMDTCGMS--TVVSTKTIKTSHYLTP 1272
DB 1386 STLSLSASNMQSGVDNSASQASST-STSTSTSESDSQSTSTSVTSQSTSESTSTSTS 1444
QY 1273 QSNHLSLKSNSVSLVPPGSSHT---LPRRAQSLKAPSLATIKSLADCNFSYTSRDA-- 1327
DB 1445 LSDSTSIKSTQS-----GSTSTASLGSSESDSQSISTSTSESKSESTSTSLSDSTS 1500

QY 1179 GENDLKFTNFGTENHRENTGRRLVBSSTSSHHKIRSQSFNTDTTSG-----ISSMSS 1234
 Db 1342 GSNLYKPCACHGDMQSGSGESGIVSEG-----DTETTINSEWCLLNAVGD 1388
 QY 1235 SPSRETVGVDAITMDTCGSMSTVSTTKITXSHVLTQSNHLSKSNVSLVPPGSSH 1294
 Db 1389 SPS-----NLETHLDPQMGDAVNVLKQK-----FTDEGSIKLPNSQSSISPGVGCN 1437
 QY 1295 TLPRAOSLKAISIATI-KSLADC-----NFSYTSRRDAFYATLKRLOQ 1338
 Db 1438 G-----KVGDLNLSITKHTPCDLGELQKHVDVFTFYDYSLQGSKLKLPIMKQSQS 1489
 QY 1339 QRMH-----PSL-----SHSEALASPAKOVLFDTTITMKANSF 1371
 Db 1490 EKAHVEDPLLRGFYFKDKSCKSKHQTTLEPDVPPHERILASHEM-----DRISKSGNI 1546
 QY 1372 ESRLTPSRFMKALSYASLDKE-DLLSP-----INQNTLQSSSVRSVSSATYG 1419
 Db 1547 EKTITGQNAKQLSHSSSIESLSPGGDLFGILGIFKNKGSDELQRTSLESWLTS--YK 1604
 QY 1420 GSDDYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGLPSTGGLVKNS 1479
 Db 1605 SNEDLFS-----CHSSGDISVSSSVGE-- 1627
 QY 1480 FHLRQMSLTEMNSIHSASLPLESTEDTGLQHTDDNCLYCVCIILGFQPSNQ--- 1536
 Db 1628 --LSKRTLDLNLRLNIQS-----TGQMSLDIASS-----INEDSAASITELSS 1694
 QY 1537 --LSAICSHSDPDIPIYDWCQEOTIHPLEVPVPSKFGISGSCSGVQSGSASTK----- 1590
 Db 1652 RVSDDITLQSSSQKSP-----TGQMSLDIASS-----INEDSAASITELSS 1694
 QY 1591 -----STELLGVKTIPTDTPMCRILLRKEVLRV-----INLSSSVSTKCHETGLTI 1639
 Db 1695 DELSCSEDIVLHKNKIPESNAS-----FRKRLTSVADESDVNVSMIVNSC--TSACTD 1748
 QY 1640 KEKYPQTFDDICLYSEVSHLLSHCTFRL 1667
 Db 1749 DE-----DD-----SLLSSSTLTL 1763

RESULT 7

US-10-172-502-4
 ; Sequence 4, Application US/10172502
 ; Publication No. US2003018583A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FOSTER, Timothy et al.
 ; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
 ; FILE REFERENCE: P07263U01/BAS
 ; CURRENT APPLICATION NUMBER: US/10/172,502
 ; CURRENT FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: US 60/298,098
 ; PRIOR FILING DATE: 2001-06-15
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 2283
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-10-172-502-4

Query Match 2.4%; Score 211; DB 14; Length 2283;
 Best Local Similarity 22.1%; Pred. No. 6.1e-05;
 Matches 156; Conservative 103; Mismatches 272; Indels 174; Gaps 25;
 QY 1016 SSIPSTLSLSESTSS-RSHNSESSEVPSSFILEDDRFSGSSTSTFFLDINEDTEPTFYD 1074
 Db 1234 STGESVSSSTSESTSLSDSTSESGTSTS--LSNSTSGSTSTST-STSISESTST--TF-- 1286
 QY 1075 RSGPIKDKNSFPFASAKLVKNRLNSLTLPNKKHRRSSDPKGGKLSSESKTSNR----- 1129
 Db 1287 -----KSESVETLSMTSTSTSLSDSTSLSDSTSDSKSDSLSTMTSTSDSISTSK 1339

QY 1130 --RIRTLTEPSVDNFHSD-----DFTPIIS-TVQKTLQLETSPWGNKHIEDTGST 1175
 Db 1340 SDSISTSTSLSGSTSESESDSTSSSEKSDSTSMISMSQSTSTSTSTSTSLSDSTST 1399
 QY 1176 -----PSIGENDLKFTKFGTE--NHNRENTSRRLVVE 1206
 Db 1400 SLSLASLMNOSGVDSNGASQASNSTSTSTSESDSTSTSTSTSTSTSTSTSTSTSLSD 1459
 QY 1207 SSTSSHMKIRSQSFNTDTTSSSMSSSPRETVGVDAITMDTCGSMSTVSTTKITK 1266
 Db 1460 STSLSKSTSGSVSTASLSGSESDSOSISTASESTS--ESASTSLSDSTSTSTNSGS 1518
 QY 1267 SHVLTPOSNHLSLSKS--NSVSLVPPGSSHTLPRAOQLKAPSIATTKSLADCNFSVTSS 1324
 Db 1519 ASTSTSLSNASASESDLSSTSL-----SDSTASMSQSSSDSTASLSLSDSTSTSN 1573
 QY 1325 RDAPGYATLKRLL-----QQORMHPSLSHSEALASPAKOVLFDTTITMKANS----- 1370
 Db 1574 R-----MSTIASLSTSVSTSESGSTSESTSESDSTSTSLSDSTSTSRSTASGSASTSTST 1629
 QY 1371 FESRLTPSRFMKALSYASLQKEDL--LSPINQNTLQSSSVRSVSSATYGSDDYIGLAL 1429
 Db 1630 SDRSTASASTSTSMRTSTSDSQMSLSTSTSTMSDSTSLSDSVSDST----- 1677
 QY 1430 PVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGLPSTGGLVKNSPHL--LRQQ 1486
 Db 1678 -----SDSTASSTSGSMVSLSDSTSTST 1702
 QY 1487 MSUTEINMNSIHSASLPLESTED-----TGLQHTDDNCLYCVCIILG 1530
 Db 1703 TSASEVMSASISDSQSMSESVDSSESVSESNESDSKSMGSTSVSDGSL-SVSTSLRK 1761
 QY 1531 FQPSNQLSAI-CSHSQFDQIPYDWCQEOTIHPLEVPVPSKFGISGSCSGVQSGSAST 1589
 Db 1762 SEVSSESSSLSCSQMSDSVSTSDSSLSVSTSLRSES-VSESDSLSDSKSTSGSTST 1820
 QY 1590 KSTEL-----LLGVKTIPTDTPMCRILLRKEVLRVNLNLSVST 1629
 Db 1821 TSGSLSTSLSGSESVSEST-----LSDSISMSDSTST 1855

RESULT 8

US-10-205-823-419
 ; Sequence 419, Application US/10205823
 ; Publication No. US20030108963A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Gannavarapu, Manjula
 ; APPLICANT: Gorbacheva, Bella
 ; APPLICANT: Hoersch, Sebastian
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Wonsay, Angela M.
 ; APPLICANT: Glatt, Karen
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Anderson, Dustin
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
 ; FILE REFERENCE: MRI-044
 ; CURRENT APPLICATION NUMBER: US/10/205,823
 ; CURRENT FILING DATE: 2002-07-25
 ; PRIOR APPLICATION NUMBER: 60/307,982
 ; PRIOR FILING DATE: 2001-07-25
 ; PRIOR APPLICATION NUMBER: 60/314,356
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/325,020
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: 60/341,746
 ; PRIOR FILING DATE: 2001-12-12
 ; PRIOR APPLICATION NUMBER: 60/362,158

;; PRIOR FILING DATE: 2002-03-05

;; NUMBER OF SEQ ID NOS: 455

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 419

;; LENGTH: 1979

;; TYPE: PRT

;; ORGANISM: Homo sapiens

US-10-205-823-419

Query Match 2.3%; Score 204.5; DB 14; Length 1979;

Best Local Similarity 17.8%; Pred. No. 0.00014;

Matches 313; Conservative 321; Mismatches 735; Indels 391; Gaps 73;

QY 51 VSNMKGHLNN-FYKLCDDIGH-----SEKLGPHVEDIIICLRALLNAEAEVRAA 102
DB 168 ISSQOEINRLSEVSRSEVGHIAQTSKAOQTDSQSEICLQNIKELKQNRSQ 227
QY 103 GLRALRYIQDSIIQKVLKLVYLIARCDIQOSNEVERTQALRLVRKMITVNASLPP 162
DB 228 EIDDHQH---EMSVLQNAHQQLKTEISRRHRE-ELSDYEERIEEL----- 268
QY 163 SVVTSNLAVNGDLQOERDRMWRACIAICELALQNPVVALRGGLNLTILKNVIDCQLSR 222
DB 269 ---ENLLQGGSGVIETD-----LSKIYEMQ-KTIQVLIQIEKVESTKQMEQLEDKID 317
QY 223 INEALITT-----ILHLLNHPKTRQVRADEVLEIRILAPYTDPHVRS---PDTAEGQLK 274
DB 318 INKLSSAENDRIIR-----RQEQNLNVEKQIMECENLKLECSKLQPSAVKQSDT 370
QY 275 EDREARFLASKGIIATFRSMAGIINLCKPGNSGIQSLIG-----VLCIPNMEIRRGLE 329
DB 371 MTEKERILASASVEEVR-----LQQAISDAENBIMRLSSLNQNSLAE 415
QY 330 VLYDFRPLPLPVTEFTFALLSVDPGRQDS-WRLSDGFVAEAKTILP-----HRA 381
DB 416 ---DNLKLRVIEVLEKESLSQSEKELQMSLLKLNNEVEIKTATRDISLDELHDL 472
QY 382 RSRPDLMDNVALIALISAFIRNGLLEGLVEVITNSD--DHSVRAVILGELL-----HMAN 435
DB 473 RLNLKAEQELNQTSIE-----KETLIAEIELDRQOEATKMLIKQOLSKQONEGD 526
QY 436 TILP-----HSHSHHLCL-PTLMMAASFDIPKEKRLASAANCLKRFHEMKRGPK 488
DB 527 SIISKQLQDNLDEKRVHQLEDDKMDITKELDVQKEKLIQSEVALNDL---HLTKQK--- 580
QY 489 PYSLHLDHIIQKAI-ATHOKRQOYLVRQXDIILKD---TEEALLINLRDSQVLQKENL 544
DB 581 ---LEDKVENLVQNLKQSNVSIQKENLELKEHTRQNEEELSRHNE--LMQSLNQ 633
QY 545 ENWNWNLIGTILKWPNNVNLRYKD---BQLHRFVRRLLYFYKPSKLYANLDDLPARAK 599
DB 634 DSNNSFKOTLLKEREAEVRLNQLNSELEQLNENLKVAF-----DVKVEN 679
QY 600 QLTVVGCQFTFLLESEEDQGYLEDVLDKIVQWLNASG-MKPSRLQNNGLLITLQSH 658
DB 680 EKLVLACEDVRHQLECLAGNQLSLEKNTIVETLKMEKEGEIEAEELCWAKKELLEANKY 739
QY 659 YFLFTGLTSLCHPHGVKMLEKCSVFQCLNLCSLKNQDHLKLTVSSLDYSRDLARVILS 718
DB 740 -----EKTIEELSARNLNTSALQLEHEHLIK-----LNQKDMETAEELCK 780
QY 719 KILTAATD-----ACLYATKHLRLVLRANVEFFNNWGIELAVTQLHDKNKTIS--- 767
DB 781 NTEQMDTDHKEYDVKVLSSELEQKQLTQINK-----EIFEKLKRSKSLQLEEL 831
QY 768 ---SEAL---DILDEACEDKANHLAIQMPALSHGDKGLILLIRF-----LSIPKG 814
DB 832 DKYSQALRKNEILRQITIEKD--RSLGSMKEENHLQEE--LERLFEEQSRAPVADPRT 887
QY 815 FSVLNERGVAKOLE--KWHREYNSKYVDLIEE-----OLNEALTYYRKVPDGNVYR 865
DB 888 LQSVTELASEVSQANTIKHEEBEIKHHQKIIEIQNQSKQLLQSLQEQKKEDEFPY--- 945

QY 866 RSNQRLQRPVYLPILHYLGQLVH-HKTGCHLLEVNITILCRNVRTPDLKWEIEIKKL 924
DB 946 -QHEQMNAATHQLFLEXDEBIKSLQKT---IEQIKTQHEERQDIQTDNSDFQETQVS 1001
QY 925 ASLWALGNIGSSNWGLNLQENVIPDILKLAKECEVLSTRGTCVYVGLIATKQGCDDI 984
DB 1002 LNI-----ENGSEKHDLKSAETERLVKIGERELEIKLLNEKN-----ISLTKQIDQL 1049
QY 985 LKCHNWDVHRSHKRLWPVVPDDVEQLCNELSSIPSTLSLNSESTSRHNSSESVSPSSM 1044
DB 1050 SK-----DEV-----GKLTQIIQKDLLEIQALHARISTSHT-QDVV 1085
QY 1045 FILEDDRFSGSSTFFLDINEDTEPTYDRSGPIKDKNSPPFPASSKLKVRILNSLTL 1104
DB 1086 YLQQLQAYAMEREREKVPFVILNEKTRENSH-----LKTEYHKMWDIVAAKEAALIKLOD 1138
QY 1105 PNKHRSSDPKGGKLSSEKTSNRRTITLTPSVDFNHSDDFTPISTVQKTL--QLETS 1162
DB 1139 ENKKLSTRFESSGQDMFRE--TIQNLRIITREKDIID-----ALSQKCTLLAVLQTS 1190
QY 1163 FMGNKHIEDTGSTPS-----IGEND-----LKFTKNFGTGHRENTSR 1200
DB 1191 STGN-----EAGVNSHOFEEELLQERDKLKQVKQWEKQVMTTVQNMQHESAQLEEL 1246
QY 1201 ERL-----VVESTSTSHMKI-----RSQFNVTDTTSGISSMSSSPSRRTVGVDATMTDC 1252
DB 1247 HOLQAQVLVDSNNKSLQVDYTGLIQSYEQNETKKNFGQELAQVQHSIGOLCNTKDILL 1306
QY 1253 GSMSTVSTTKITKTSHTTPQSNH-LSLSKNSVSLVPPGSSHTLPRAQSLK---APSI 1308
DB 1307 GKLDII--SPOLSSASLLTPQSAECLRASKSEVLUS-----ESSELLQLEELERKSLQEK 1360
QY 1309 ATIKSLADCNFSYSSRDAGYATLKRIQQOORHPSLSHSEALASPAKDVLTDTITMKA 1368
DB 1361 ATIRTLQENHRLSDSTAA--TSELERKEHQTDSIEIKQLKEKQDVQLKLEKDLIIKA 1418
QY 1369 NSFESRLTPSRFMKALSASLDKEDLLSPINQTLQRS--SSVRSMVSSATYGGSDDIIG 1426
DB 1419 KS--DQLLSS-----NENFTKNVENELLQAVTNLKERI----- 1451
QY 1427 LALPVDIN-----DIFQVKDIPY-----FOTKNIIPHD-DRGARAPA 1462
DB 1452 LILEMDIGKLGKENEKIVETIRGKETETYOALQETNMKFSMMLREKEFECHSKEKALAFE 1511
QY 1463 HDAGLPGSGTGLVKNKSFHLRQOMSLTEIMNSIHSASLFL--ESTEDTGLQEHDDNC 1520
DB 1512 QLLKEKEQKGTGELNQLLNNAVSKMQEKTIVVFQQRDQVMLALKQKQMENTALQNEVQR-- 1569
QY 1521 LYVCIEILGFQPNQLSAICSH-----SDFQDIPYSDWCBOETHNPLEVVPKFSGIS 1574
DB 1570 ----LRDKFPRNQELERLNRHLLSESDSYTREALAAEDREAKLRKKVTVLEEKLVSS 1624
QY 1575 GSDGVSGEGSASSTKSTELLGVKTIPTDTPMCRILLRKEVLRLVINLSSSVSTKCHET 1634
DB 1625 NAMENASHQASQVQVESLQEQNLNVSKQDETALQVSQEQVKQYALSAN----- 1675
QY 1635 GLITIKYKYPOTFDDICLYS 1654
DB 1676 -LQWVLEHFQ--BEKAMYS 1692

RESULT 9

US-10-437-963-137285

;; Sequence 137285, Application US/10437963

;; Publication No. US20040123343A1

;; GENERAL INFORMATION:

;; APPLICANT: La Rosa, Thomas J.

;; APPLICANT: Kovalic, David K.

;; APPLICANT: Zhou, Yihua

;; APPLICANT: Cao, Yongwei

;; APPLICANT: Wu, Wei

;; APPLICANT: Boukharov, Andrey A.

;; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 137285
 ; LENGTH: 2052
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(2052)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_38782C.1.pap
 US-10-437-963-137285

Query Match 2.3%; Score 204.5; DB 16; Length 2052;
 Best Local Similarity 18.9%; Pred. No. 0.00015;
 Matches 318; Conservative 266; Mismatches 605; Indels 495; Gaps 79;
 Qy 51 VSNRKLGLHNNFKLL-----CDI-GHSEKLGK-----HYEDII----- 86
 Db 353 LENTRKSEL-SPTPLVAADPCDLAGSPKEMAFSDPANQNDLMIVPKVPTGVSGTALAA 411
 Qy 87 -CLR-----LALLNEAKEVRAAGLALRYLIQDS-SILQKVLKLVVDVLIARCIQDSNE 140
 Db 412 SLLRDGSLAKIYKYQATDARHERFGRHAFAVLRLVH-----BIEKAE 460
 Qy 141 V---ERTQALRLVRMTVNASL-----FPSSVTN-----SLIANGNDG 176
 Db 461 LILDERAHRMRVAYALMDQQLLEHDFNTIRNLKSELKRERDHSIAQKID 520
 Qy 177 LQERDM-----VRACIAIIC-----ELALQNEPVVALRGGLN 209
 Db 521 LQKIKLSIYLTAPTAQACILGVCTSKNYIDINGHCGAVLKIAVLLKCEQDIQLRCG-- 578
 Qy 210 TILKNVDCQLS-----RINE-----ALITLILHLN--HPKTYQVVRAD 247
 Db 579 SSLPVGDLSTSTGVPVENNIHEMTFNDINGLVQNVQLRQVHLLSADLDKRD 638
 Qy 248 VELERI-----LAPYTDHYRSPDTAEQLKEDREARFLASKMGIIATFRSWAGIINLCK 303
 Db 639 MELRESQIELKKITD---DAASRVEKVMKSEEQAIM-----IESLRSVAMYKLCSE 689
 Qy 304 PGNISQISGLVCLIP-NMEIRGLLEVLVDIFRLPLPVVTEEFIEALLSVDPGRFQDSW 362
 Db 690 EQQ---KSRSNVEHIPKNLE-----DGRKDLMLVFE-----GSQEVTS 724
 Qy 363 RLSDFVAAAEAKTILPHRARSPPDLM---DNYLALILSAFIR---NGLLEGLVEVITNS 415
 Db 725 RKAYEQVSERAKSLDEELTKURLLSURSERDKAVLEAEFARERLNGFTABL--EHORKE 783
 Qy 416 DDHISVRATILGELLHMANTILPHSSHHL---HCLPTLMNMAASPDIPKEKRLASAA 472
 Db 784 ANSISLRAELMHLVDYKEKRENSSESMKAVEENSKLMMESI---LKNKEKILAKSE 840
 Qy 473 LNCILKRHEMKRGPKPY-----SLHLDHI IQKAIATHOKRDQ-----YLRVOKDI 518
 Db 841 KKALEEVHDLTTRVHRLQATDITTHATEVQENARSIERRQNEYIKRLERDMAEVKEL 900
 Qy 519 FILKTEALLINRDS-----OVLQHENLEWNL-----IGTIL 555
 Db 901 QEQRDHVRVLTLDKKNAPDGLCKLVQDMRKELQNSKAAATDAESRAAAVAEAKCSDLETKL 960
 Qy 556 KWPNNV-----LRNYKDEQLHRFVRRLLYFYKPSKLYANLDDFAKAKQLTVVGC 606
 Db 961 KSRKENDELFOLEKELEYKEE-----AQANKSYM-----L 991
 Qy 607 QFTBFLLESEBDGQYLEDLVKDIQVOMLNASSGMKPSRQNNGLLTLTSHYFLFTGL 666

Db 992 QYKE-IANSNESALKQMESALQD---FKTESETIKKLSLEDEITKLRTKISE----- 1038
 Qy 667 SCHPHGVKMLEKESVFOC---LLMLCSLKNODHLLKLTVSSLDYSDRGDLARVILSKILTA 724
 Db 1039 -----LEKCYIMKCEBAASAIAEKD-----TTSIMKEISVLRNESEKVIQIE 1083
 Qy 725 TDACRLYATHKHLRVLRLANVEFNNWGIELLVTLQHDKNKTITSEALDIDDEACEDKANL 784
 Db 1084 KLETELASSK-----RALDEQYKRW---RSAQDNVERQVILQSETTQELTTSKLSL 1134
 Qy 785 HALIQM-----KPAISHLGDKGLLLALLRFLSIPKG-----FVSLNERGYVAK 826
 Db 1135 QOEIIVLRQVETQKAEENDGLRTLGBOEKIELV-----KGDEALQYELANDQNKIL- 1187
 Qy 827 QLEKWHREYNSKYVDLIEEQLENEA-LTTYRKVPDGDN-----YVRSNRQRLQRPVYL 878
 Db 1188 -----HNQLEALHRLAEKERNIAGLSQRIIDSHGEDDLHSVIGYLRRSKEILE----- 1236
 Qy 879 PIHLYGOLVHHKGTGCHLLEVNIIITELCRNVRTPDLDKWEIIEKKLAKSL--WALGNIGSS 936
 Db 1237 -----SALKSTK-----EAQDLRSQTDARSATML--KDEEFKSLQQLVFFCIKNALPQ 1284
 Qy 937 NWGLNLQEEVNV-----IPDILKLAKQCEVLSIRGTCVVYVGL-----IAKTK- 979
 Db 1285 VRELNLRESNIQLREENKINFECQKFHDQAKAKMEARLHNLLEKQVDAEICKKEI 1344
 Qy 980 --QGCDITLKCHN--WDAVHRSRKHLWVPVDDVQELCNELSSIPSTLSLASE--STSSRH 1033
 Db 1345 EMOKTEIANLNQKISELVENS---GVDLSTVETMKDELONIKATILRENSABLERPKL 1400
 Qy 1034 NSESEVSPSMFILEDDRFSGSSTSTFFLDINEDTEPTFYDRSGPIKDKNSPPFFASSKL 1093
 Db 1401 LSEKDSVIRN---LEEKLAGCQSE---LDARE-----KKLNDVEASLKSEI 1440
 Qy 1094 VKRILNLSLTPNKHRSRSDPKGKLSSEKSTSNRI-----RTLTPSVDFPHSD 1145
 Db 1441 DRHKIN-INIKRKLDAKAKE--ELTKEQSLSKQLEDLKKXGOKTITTESENQAIKE 1497
 Qy 1146 DFTPISTVQKTLQETSFMGNKHIEDTGSTPSIGENDLKFTKNFTGTEHNR-----ENTS 1199
 Db 1498 KDFRIQTLKVLKE-----RDNKKEKAFRRRNEKVFYTAIQMN 1538
 Qy 1200 RELVVESSSTSHMKIRSQSFNTDTTTS-----GISSMSSPS----- 1237
 Db 1539 QERKQVEESTEKHRQAVKEVIEHYVTGISSQIPSGSAIDEQLRSVFLAIKAVEESPSPQD 1598
 Qy 1238 ---RETGVDAATTMDTDCGSMSTVSTVTKITKTHYLTPQSNHLSLSKSNVSLVP--PGSS 1293
 Db 1599 GAASQTPSVESADVDASAATAGQVAT-----PPRPAQVKVVEERAVSTLPKPESTE 1649
 Qy 1294 HTLPRAQSLKAPSIATIKS-LADCNFSYTSRDAFGYATLKRLQ---QRMHPSLSHSE 1349
 Db 1650 VRPGRPLVRPSLREVEEPQADIDTTVVEGSTEGKGLMERETPGCVGALQPS--SRKR 1708
 Qy 1350 ALASP--AKDVLPTDTITMKANSFESRLTPSRFMKALSYASLDKEDLLSPI----- 1398
 Db 1709 LIFSPTQDDASQGETTDANPPLKPKGSSQCTSELKTEQSPHEDVMAVPVLPSTELD 1768
 Qy 1399 -----NONTLQSSSVRSVMVSSA--TYGGSDD---YIGLALPVD-----INDIFQV 1439
 Db 1769 EQQFGEEMDTQSSLPFVEEETREDLDGDKDMETHIDASMDIQOGDAETGIDNDATT 1828
 Qy 1440 KDIP 1443
 Db 1829 EDVP 1832

RESULT 10
 US-10-282-122A-71033
 ; Sequence 71033, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 71033
LENGTH: 1831
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1801)..(1801)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1805)..(1805)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1808)..(1808)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1809)..(1809)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1825)..(1825)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1828)..(1828)
OTHER INFORMATION: X=any amino acid
US-10-282-122A-71033

Query Match 2.3%; Score 200.5; DB 15; Length 1831;
Best Local Similarity 23.1%; Pred. No. 0.00024;
Matches 129; Conservative 74; Mismatches 256; Indels 99; Gaps 17;

QY 1013 NELSSIPSTLSINSESTS-----SRHNSESESVPSMFILEDDRFSGSSTSTFFLDLINED 1067

Db 1271 SESSSTSKKLS-ESASTSMDSASASTSESNSTSL-----SGSTSTSLSGSTST 1320
QY 1068 TEPTFYDRSGPIKDKNPPFPASSKLVKNRILNLTLPNKKHRSDDPKGKLS-----S 1122
Db 1321 STSESASTSTSVSDSNS-----ASTSLSES---TSTSLSDSTSTSDSTASTSESDSNS 1373
QY 1123 ESKTSNRRIRLTLPESVDNFHSDFTPISTVQKTLQLETSPMGKHIEDTGSTPISGEN 1182
Db 1374 TSTSMSESLSTSVSDSTSTSDSTASTSVSDSNSASTS-----LSDSSTST-SLSDST 1426
QY 1183 LKFTKNFGTE--NHRENTSRRLVVESSSTSHMKIRSQSFNTDTTSGISSMSSSPSRET 1240
Db 1427 SASTSDSASESASESTSESTSVSESSST-----SVSDSTSTSEASTSESEST 1480
QY 1241 -----VGVDATTMDTCGSMSTVSTKTIKTSHYITPQSNHLSLKSNSVS----- 1286
Db 1481 SESTSVSESSSTSVSDSTSTSTSEASTSESTSESTSLSGSSTSVSDSTSTSE 1540
QY 1287 -----LVPPGSSHTLPRAQSLKAPSTATIKSLADCFSTSSRDAPGY 1330
Db 1541 SASTSTSVSDSNSSTSTSESTSTSLSNSSTSTSDSASTSTSVSDSNSASTSLSGSL-- 1598
QY 1331 ATLKLQOQRMHPSLSHSEALASPAKDVLFDTITMKANSFESRLTPSRFMKALSYASLD 1390
Db 1599 -----STNVSDSTSTSDSASTSTSESDSNSASTSLSGSLSTSDSTST 1645
QY 1391 KEDLLSPINQNTLQRSSVRSMVSSATYGGDDYIGLALPVDINDIFQVKDIPYQTKNI 1450
Db 1646 TDSASTSTSEASTSTSVSESDSESTSMVSDSNSSTSTSLSDSTSTSVSDSTRTST 1703
QY 1451 PHDDRGARAFADAGLPGSGGLVKNFSF---HLLROOMSLTEIMNS---IHSDASLFL 1504
Db 1704 ---DSASTSKVSESNASTSVSGSTSTISDSSSTSTSMSTSETFTSQSPINSESQFIG 1760
QY 1505 ES-TEDTGLQEHDDNCL 1521
Db 1761 DLSLSDTTVTSQKNTNML 1778

RESULT 11

US-10-470-048B-268
; Sequence 268, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE OF INVENTION: ANTIGENS TO A SPECIFIC PATHOGEN
; FILE REFERENCE: SONN:035US
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 268
; LENGTH: 1870
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-470-048B-268

Query Match 2.3%; Score 200.5; DB 17; Length 1870;
Best Local Similarity 23.1%; Pred. NO. 0.00025;
Matches 129; Conservative 74; Mismatches 256; Indels 99; Gaps 17;

QY 1013 NELSSIPSTLSINSESTS-----SRHNSESESVPSMFILEDDRFSGSSTSTFFLDLINED 1067
Db 1271 SESSSTSKKLS-ESASTSMDSASASTSESNSTSL-----SGSTSTSLSGSTST 1320
QY 1068 TEPTFYDRSGPIKDKNPPFPASSKLVKNRILNLTLPNKKHRSDDPKGKLS-----S 1122
Db 1321 STSESASTSTSVSDSNS-----ASTSLSES---TSTSLSDSTSTSDSTASTSESDSNS 1373
QY 1123 ESKTSNRRIRLTLPESVDNFHSDFTPISTVQKTLQLETSPMGKHIEDTGSTPISGEN 1182

Db 1374 TSTSMSESLSTVSDSTSTSTSDSASTSTSVSDNSASTS-----LSDSTST-SISDST 1426
Qy 1183 LKFTKNFTE--NHRENTSRRLVVESTSHMKIRSQSFNTDTTSGISMSSSPSRET 1240
Db 1427 SASTSDSASASASESTSESTSVSEST-----SVSDSTSTSESASTSTSEST 1480
Qy 1241 -----VGVDATTMTDCGSMSTVSTKTIKTSHYLTQSNHLSKSNVS----- 1286
Db 1481 SESTSVSESTSVSDSTSTSTSESASTSTSESTSESTSESTSESTSESTSESTSE 1540
Qy 1287 -----LVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSTSRDAFGY 1330
Db 1541 SASTSTSVSDNSNSTSLSESTSLNSNSTSTSDSASTSTSVSDNSASTSLGSL-- 1598
Qy 1331 ATLRLAQOQRMHPSLHSEALASPAKDVLTDTITMKANSFESRLTPSRFMAKLSYASLD 1390
Db 1599 -----STNVSDSTSTSDSASTSTSESDNSASTSLGSLSTISDSSTST 1645
Qy 1391 KEDLLSPINQTLQRSSVRSMWSAYVGGDDYIGLALPVDINDIFQVKDIPYFQTKNI 1450
Db 1646 TSDSASTSTSESASTSTSVSESDSESTSMVSVDNSNSTSLSDSTSTSVSDSTSTST 1703
Qy 1451 PPHDRGARAFADAGGLPSTGTGLVKNF--HLLRQOMSLTEIMNS---IHSDSLFL 1504
Db 1704 ---DSASTSVSESNASTSVSGSTSTSDSSTSTSTSTSTSTSTSTSTSTSTST 1760
Qy 1505 ES-TEDTGLQHTDDNCL 1521
Db 1761 DLSLEDITVQSKNTNML 1778

RESULT 12
US-10-369-493-5784
; Sequence 5784, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5784
; LENGTH: 3507
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5784

Query Match 2.2%, Score 197; DB 15; Length 3507;
Best Local Similarity 19.3%; Pred. No. 0.0012;
Matches 186; Conservative 138; Mismatches 335; Indels 304; Gaps 42;
Qy 902 ITELNRVTPDLDKWEIKKLKASLWALGNIGSSNGLNLQENVPDILKLAKQCEV 961
Db 1802 IEPEPKSDKTACTDEWSRLCELEKKQCTV-----DEERV-----CCGA 1840
Qy 962 L-----SIRGTC--VYVLGIATKQGC-----DILKCHNWDVVRSHKHLWP----- 1002
Db 1841 CLPQHPHPIGTQSLQISGLCAQ-KNCKNKAECIDI-----HPDSHFCSPPDG 1889
Qy 1003 -----VVPDDVEQLCN-----ELSSIPSTL-SLN----- 1025
Db 1890 IGDGMICDDVDE-CNNAGMCDDEWTKCENTIGSFNCVCLSGPKVKVDEKVKQPNRE 1948
Qy 1026 -----SESTSRHNSSESVSSMPILEDRRGSSSTSTFFLDINEDTEPTFYDRSGPIK 1081

Db 1949 KIBIDENSSSSGQEK-PTTKGIVSSTSATSEST-----AEPHV---TTSISS 1996
Qy 1082 KNFPFPFASSKLVKNRLNS---LTLPNKHHSSSDPKGKGLSSSEKTSNRRI----- 1131
Db 1997 TTKTQMTSSKSPENVTMSSESPEVSTSSKSTASSETTSSVSTPSSSSSEAPLTSSPAT 2056
Qy 1132 --RTLTPSPVDFNHSDDFTP-----ISTVQKTLQLETSFMGNKHIEDTGSTPS----- 1177
Db 2057 TTEVITESSV-----KSTTPKBESSSEITVKLSKSPKSEVTESSVKSSPSTTSSQSVTS 2111
Qy 1178 -----IGENDLKFTKNFGTENHRENTSRRLVVES-----TSSHMKIRSQ 1218
Db 2112 TVPETSSTVLSSEAPVSTSTPTEVSTSETKPSLSSASTTGTNTSTPSTSSLASVKST 2171
Qy 1219 SFNTDTTTS-----GISMS---SSPGRSTV-GVDATTMTDCGSMSTVWSTKTIKTSHYL 1270
Db 2172 SAPEGTSASVAPVKLSLSPDVQSPSTKTPTDATESSTVQASETSSGTSVKSTSEPESHV- 2230
Qy 1271 TPOSNHLSSKSNVSVLVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSTSRDAFGY 1330
Db 2231 ---TKLSITSSNPPSSVPVTPSPKSTPTVPESTEQTSTTPSG----- 2269
Qy 1331 ATLRLAQOQRMHPSLHSEALASPAKDVLTDTITMKANSFESRLTPSRPM-----K 1382
Db 2270 -----QSLTPMNSNSELTTSEPHVL---SSLSPOVVSQSSSTTPNNLSESTVETPK 2318
Qy 1383 ALSVASLDKED-----LLSP---INONTLQRSSV---RSMVSSAT----- 1417
Db 2319 TSSEVLSNSEPSTTEAPTTLSPDILSTNNLSQSTSTVSTEDRSEISSENSEKPTSAPE 2378
Qy 1418 -YGGSDYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGAFAHADAGLPGTGGLV 1476
Db 2379 LVTSSVTHVASSSPDVPTSESEBDDLTGSSSTENIP---EASSKQTSSTPTDPTTASEE 2435
Qy 1477 KNSFHLLRQOMSLTEIMNSHSDASLPESTETDGLQHTDDNCLYCVCIILGFQPSNQ 1536
Db 2436 PTKSTSMSPDLSTT---SNVLESSTTPSSSKSPVSSSTEG----- 2474
Qy 1537 LSAICSHSPDIPYSDWCETIHNPLE-----VWPS---KFSIGSCSDGVSQSGAS 1587
Db 2475 ISVVTST-TEFSKVP-----ESTISSVLEEDLTKTFTSPILEETTASSETSEPLETDSLTV 2528
Qy 1588 STKSTELLGVKTIPTDTPMCRILLRKEVLRLVINLSSSVSTKCHE--TGLLTIKEKYPQ 1645
Db 2529 SVRIHELTTSENVPKES------TTTSSSSKPSQBPAGILSTVVVPT 2574
Qy 1646 TFDDICLYSEVSHLSHCTFRLPCRFIQLFQDVQFLOMHEAEAVLATPPKQIVDTS 1705
Db 2575 SSVSLITASEIAITNTPFK-----QORTPITTSKSLVKSTT 2613
Qy 1706 AES 1708
Db 2614 SPS 2616

RESULT 13
US-10-245-802-18
; Sequence 18, Application US/10245802
; Publication No. US20030124134A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, John E.
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND METHODS TO VACCINATE AGAINST
; FILE REFERENCE: 013361.4003
; CURRENT APPLICATION NUMBER: US/10/245,802
; CURRENT FILING DATE: 2002-09-13
; PRIOR FILING DATE: 2000-11-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1443
; TYPE: PRT

```
; ORGANISM: Candida albicans
US-10-245-802-18

Query Match      2.2%; Score 195.5; DB 14; Length 1443;
Best Local Similarity 23.2%; Pred. No. 0.00037;
Matches 171; Conservative 107; Mismatches 295; Indels 163; Gaps 33;

QY 1014 ELSS---IPST---LSLNSBSTSRHNSSESVSPSMFLEDDRFSGSSTSTPFLDINE- 1066
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 695 ELSRSDIPSSIGLSSSESTVSSVSDSVSSSTSESIASSYDSYSSSIEBSTLSSDR 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1067 -----DTEFTFYDRSGPIKOKNSPPFFA-----SSKLKVRNLNLSLTLPNKKHRSDDP 1115
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 755 CSSSISDT-TSFYDSSSDLESTSYSSSIDAQSHLVQS-VSNSISITSQELSSSSEE 812
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1116 KG-----GKLSSE-----SKTSNRIRLTETPSVDFNHS-----DFTPISTVQKTLQLE 1160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 813 SSTFATLVSSDASSILSDSTSYSPSTIGSSDDPFPHTIAGESDLSISFITSTVEIS 872
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1161 -----TSFMGNKH-EDTGTPSIGENDLKFTKNFGT----- 1191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 873 SDSVSLTSPASSFDDSSSLNDSSSPSDQSDILTSSFSFTLVVPFSLSSSSLSLT 932
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1192 -ENHRENTGRRLVVESTSSHKIRSQFNFTD-----TTGIGSSMSSPS---RE 1239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 933 YPHVNSTTTHASESSSVASPMASESANDDTYLTSESDTTSIGTDSSTVTFPCRD 992
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1240 -----TVGVDATMD-----TDCGSMSTVVSXT-----IKTSHYLPQSNH 1276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 993 NGDGCIVTGWPSDDISDEQSDVTTSFVASTPTSAEQSITDNPNDSSQSSSTK 1052
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1277 LSLSKS-----NSVSLVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSTSRDAFGYAT 1332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1053 LSVSVSDTVVNSIL---SETSLSDSDTSDDTSISSTNTSDTGNVAGSHSTASIK 1109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1333 LKRLQORMHPSLSHSEALASPAKDLFTDTITMKANSFESKLTSPRFMKALSYASLDKE 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1110 ESSIQKTGTVLSSSVLSTKLSSTSDIT-TELITTELITTELITTELITTE-----DNE 1161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1393 DLLSPINQTLQSSSVRSVMSVATYGGSDDIIGLALPVDINDIPQVKDIPYFQKNIPP 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1162 -----PNTFTSPSHSEIFS-----SD-----NSVLS-KQVDCGSTEIIPP 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1453 HDRGARAFADAGGLPSGTGLVKNFSHLLRQOMSLTEIMNSIHSASLFLESTEDTGL 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1198 VTDT-TTVSSSVHSTEASTALGENSFKVASAPNTE--TSLSTSSSNHATESGOT 1254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1513 QEH-----TDDNCLVCVCIIEILGFOPSNQLSAICSHSDPQDIPYSDWCQETH 1560
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1255 VKSEASAEAIPTSTNRLSYSTEEAKGSTYPNSGNTNLMWESQVAAPDSTSVLTA 1314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 NPLEVVPKFGISGCDGVSGEGSASTKSTELLGVKTIPTDPT--PMCRILLRKEVLR 1618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1315 NP--VTSTFDKKS--SAAVNQ---PSKTSIEESIGSLDSVNETNNGFIATLSQSEAPN 1367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1619 LVINLSSSVSTKCHET 1634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1368 SLIH-SEISITWAKT 1382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-815-242-13057
; Sequence 13057, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
```

```
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13057
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13057

Query Match      2.2%; Score 193; DB 9; Length 560;
Best Local Similarity 22.3%; Pred. No. 0.00012;
Matches 139; Conservative 98; Mismatches 243; Indels 144; Gaps 24;

QY 1016 SSIPFTLSLNESTSRHNSSESVSPSMFI-----LEDDRFSGSSTSTPFLDINETEP 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 33 TSISTSTSI-SESTST---FKSESVSTSLSMSTSTSLSD---STSLSTSLSDSTSDSKS 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1071 TFYDRSGPIKOKNSPPFPASSKLKVRNLNLSLTLPNKKHRSDDPCKGKLSSEKTSNRR 1130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 DSLSTSMSTSDS-----ISTSK--SDSISTSTSLSGSTSESESDTS---SSEKSDS-- 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1131 IRTLTPSVDFNHSDDFTPISTVQKTLQETS-----FMGNKHIE-----DTGST 1175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 ----TSMISMSQSTSGSTSTSTSTSLSDSTSTSLSLASNMNQGVDSNSASQASNSTS 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1176 PSIGENDLKFTKNFGTE--NHRENTSRERLAVVESTSSHKIRSQFNFTDTTGIGSSMS 1233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 189 TSTSESDSQSTSYTSQSTSESTSTSTSLSDSTSTSKTSQSGSVSTSLSGSESES 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1234 SSPRETGVGDATTMDTDGSMSTVVSXTKITKTSHTLPQSNHLSLSKS--NSVSLVPPG 1291
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 DSQISTSASEST-ESASTSLSDSTSTNSGASASTSTSLNNSASASESDLSSTSL----- 303
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1292 SHTLPRAQSLKAPSIATIKSLADCNFSTSRDAFGYATLKRL-----QQORMHPS 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 -SDSTASMQSSESQDSQSTSLSDSLSTSTSNR---MSTIASLSTSVSTSESGSTSES 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1345 LSHSEALASPAKDLFTDTITMKANS-----FESRLTPSRFMKALSYASLDKEDL--LS 1396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 TSESDSTSTSLSDSQSTSRSTSTASGASASTSTSTSDSTSTSTSMRTSTSDSQSMELS 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1397 PINQNTLQSSSVRSVMSVATYGGSDDIIGLALPVDINDIPQVKDIPYFQKNIPPHDDR 1456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 TSTSTSMSTSDTSLSDSVSDST----- 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1457 GARAFADAGGLPSGTGLVKNFSHL---LRQOMSLTEIMNSIHSASLFLESTEDT--GL 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 -----SDSTASSTSGSMVSVISLSLSDSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1513 QEHDDNCLVCVCIIEILGFOPSNQLSAICSHSDPQDIPYSDWCQETHPLEVY--PSKF 1570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 SESNSES-----DSKMSGSTSVSDSGSLVS-----TSLRKSSESSESSEL 533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1571 SGISGCDGVSGEGSASTKSTEL 1594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

•
—
—

•
—
•
—

•
•
—

•
•
—

•
•
—

•
—
—
—
•

—
•
—

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2005, 13:48:50 ; Search time 27 Seconds
(without alignments)
4722.244 Million cell updates/sec

Title: US-10-782-244-3

Perfect score: 8811

Sequence: 1 MAAIGRSLKRLVRGRND.....AEAVLATPPKQPIVD7SAES 1708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	615	7.0	123	4	US-09-513-999C-5019
2	544	6.2	826	4	US-09-248-796A-14704
3	281	3.2	149	4	US-09-270-767-33327
4	281	3.2	149	4	US-09-270-767-48544
5	212	2.4	2137	3	US-09-134-001C-4463
6	204.5	2.3	1979	4	US-09-949-016-6468
7	204.5	2.3	2047	4	US-09-949-016-7404
8	196.5	2.2	1306	4	US-09-538-092-330
9	189	2.2	823	4	US-09-248-796A-16699
10	189	2.1	2224	4	US-09-054-272-38
11	179.5	2.0	3854	4	US-09-949-016-7876
12	176.5	2.0	2842	1	US-07-741-940-7
13	176.5	2.0	2842	1	US-08-289-548A-7
14	176.5	2.0	2842	1	US-08-452-654-7
15	176.5	2.0	2842	1	US-08-449-731-7
16	176	2.0	2843	1	US-08-452-655B-2
17	176	2.0	2843	1	US-08-452-655B-7
18	176	2.0	2843	3	US-08-450-582-2
19	176	2.0	2843	3	US-08-450-582-7
20	176	2.0	2843	4	US-09-538-092-1007
21	176	2.0	2954	4	US-09-150-867-1
22	176	2.0	2973	2	US-08-821-355A-7
23	176	2.0	2973	2	US-09-003-687A-7
24	176	2.0	2973	3	US-09-136-605-7
25	174	2.0	2843	4	US-10-092-138A-30
26	172	2.0	2568	4	US-09-866-108A-3
27	171	1.9	2662	4	US-09-595-684B-31

28 170.5 1.9 569 4 US-09-248-796A-16697 Sequence 16697, A
29 170.5 1.9 2663 4 US-09-538-092-1252 Sequence 1252, Ap
30 170 1.9 2843 1 US-07-741-940-2 Sequence 2, Appli
31 170 1.9 2843 1 US-08-289-548A-2 Sequence 2, Appli
32 170 1.9 2843 1 US-08-452-654-2 Sequence 2, Appli
33 170 1.9 2843 2 US-08-370-235A-2 Sequence 2, Appli
34 170 1.9 2843 4 US-08-449-731-2 Sequence 2, Appli
35 169.5 1.9 5024 4 US-09-710-279-2964 Sequence 2964, Ap
36 169 1.9 897 1 US-08-095-737-4 Sequence 4, Appli
37 169 1.9 897 1 US-08-480-145-4 Sequence 4, Appli
38 169 1.9 897 2 US-08-477-389-4 Sequence 4, Appli
39 167.5 1.9 10182 3 US-09-134-001C-3159 Sequence 3159, Ap
40 167 1.9 1140 4 US-09-538-092-647 Sequence 647, App
41 167 1.9 3878 4 US-09-914-259-11 Sequence 11, Appli
42 165.5 1.9 2332 1 US-08-276-594A-2 Sequence 2, Appli
43 165.5 1.9 2351 1 US-08-366-851A-2 Sequence 2, Appli
44 163.5 1.9 2332 1 US-07-864-004B-4 Sequence 4, Appli
45 163.5 1.9 2332 1 US-08-251-937A-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-513-999C-5019
; Sequence 5019, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCES: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5019
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-513-999C-5019

Query Match 7.0%; Score 615; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.7e-45;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAIGRSLKRLVRGRNDGSEENVPLDITREPSDNLREILQNVARLQGVSNRKLGLHL 60

Db 1 MAAIGRSLKRLVRGRNDGSEENVPLDITREPSDNLREILQNVARLQGVSNRKLGLHL 60

Qy 61 NNFTKLLCDIGHSEELKGFHYEDIIICLRALNKAKEVRAAGLRALRYLIQDSSIIQKV 120

Db 61 NNFTKLLCDIGHSEELKGFHYEDIIICLRALNKAKEVRAAGLRALRYLIQDSSIIQKV 120

Qy 121 LKL 123

Db 121 LKL 123

RESULT 2

US-09-248-796A-14704
; Sequence 14704, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

```

; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14704
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (826)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-14704

Query Match
  6.2%; Score 544; DB 4; Length 826;
Best Local Similarity 25.6%; Pred. No. 1.8e-37;
Matches 231; Conservative 152; Mismatches 327; Indels 194; Gaps 40;

QY 277 REARFLAKMGIIATPRSWAGIINLCKPCNGSIQSLIGVLICIPNMEIRRGLEVLVDIPR 336
Db 3 QGISFLISTL-----LKNFGIIAYLINDFTSIKLLMNMLOKNTKVRISLDLDDALM 57
QY 337 L-PLPVVTEETFEALLSVDPGRFQDSWRLSDFVAAE---AKTILPHRARSPDLMDNYL 392
Db 58 IQSPFWLKNPSIGDSIMRYNELFNDGYQPKPEFKPIDDEFKSNILHH-----QGLL 109
QY 393 ALILSAFIRNGLEGLEVI-----TNSDDHISVPAATILGELLHMANTILPHSHS 443
Db 110 TLIL---MENGWKHLQTOIEENRGIDIPRSIEH---KEATMLLNLVAMANNYLPaelI 164
QY 444 HHLHCLPTLNNMAASDIPKEKRLRASAALNCLKRFHEMKRGKPKP-YSLHDHIIQKAI 502
Db 165 GKLLLPDL-SLCAFEIENFR-----HEFR---PTETNQHVKSIIQ-- 205
QY 503 ATHQRDQYLRVQKQDIFILKOTEEALLNLRDSQVLQHKENLEWNNLIGTILKWPVNL 562
Db 206 -----LNIQKYNIDNFTFKTVSN---SKILAVKEPEDNNWQLLTILQGLGNP 253
QY 563 RNYKD--EQHFRFVRLLFYFKPSKLYANLDL--DPAKAKOLTVVQOFTFLESEED 618
Db 254 KRFDEVLKVPKPKFKLLSYRPFYRFTSVTLTKDKNSPKYN-IGCOLIELFL-SLDN 311
QY 619 GQGYLE----DLVKDIVQWLNASGM-KPERSLQNNGLTTLTSLQHYFLFICTLSCHPHG 672
Db 312 GKLYSKSLLPQISEIQAIDPFSGITKDPILSEKRLTETASGVYLRFGVLSSHSLG 371
QY 673 VMLEKCSVFQCLNLCSLKNQ---DHLKLTV-SLDYSRDGLARVILSKILTATDAC 728
Db 372 IRMLEQWQFTFLNLISSTSETNLFIITFKYADFIDSFPFNTLEMAIRVSNFKI 431
QY 729 RLYATKHL--RVLLRANVEFFNNGIELLVTQJH-----DKNKTISSEALDILDEAC 778
Db 432 RNYILKHLVPLRIKTECEKRLFT---IRFLVENLYCIGPSENGGGEIVTKSIELIYHY 488
QY 779 EDK--ANLHALIQMPALS-----HLGDGLLLRLFLSIPKPSYLNERNGVVAKOLEKW 831
Db 489 QEDNFSNLDTLIQPLVLEVKYRLGKK---LLNNFLRFPPEGFKYLERGLVLEEKFNKL 545
QY 832 HREYNSKYV-----DLIEEQ-----LNEALTYYRKPVDGNYVYRRSNQRLQRPHV 876
Db 546 TQALDPKLYKINSIIQYFYPVHFANNLNA-----DNDN-----NDADDQLSI 592
QY 877 YLPIHL-----YGQLVHHTKGCH--LLEQNIITELCRNVRTPDLKX--EIKKL- 923
Db 593 YFLKNLLSTEGNFPFGKGEFLHSILLSTEIIFQOISQAKFHDENQDDEVKFLI 652
QY 924 ---KASLWALGNISGNNWGLNL-----QEENVIPDILKLAKOCEVLSIRGTCVY 970
Db 653 NAVQNLMILGNALGEFGIQLDDPMYNNLNANNSINIILNFKTCSIQWIRGICY 712
QY 971 VLGLIAKTKQGCDDILKCHNW-----DAVRHSRKHLPVVPDDVEQLCNELS----- 1016

; Db 713 VLGMIASTIEGTEILDEFNWVSCVDQYGNCKRLSYPKVENLVEIFSIEMSNPYRDRYXH 772
; QY 1017 ---STPSTLSLNSESTSSRHNSSESVPSSMFEILDDRFSGSSTSTFFLDINDETEPTFY 1073
; Db 773 IFNSIP--VEVTEANSNTINNDTGNIEB-----LVISDNTDKPLV 811
; QY 1074 DRSG 1077
; Db 812 DENG 815

RESULT 3
US-09-270-767-33327
; Sequence 33327, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 33327
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33327

Query Match
  3.2%; Score 281; DB 4; Length 149;
Best Local Similarity 38.5%; Pred. No. 5e-16;
Matches 55; Conservative 35; Mismatches 51; Indels 2; Gaps 1;

QY 123 LKVDYLIARCDIQOSNEVERTQALRVKMITVNASLFPSSVTNSLAVNGDGLQERDR 182
Db 9 LQLOHLLCRSIDLMLKNDDERVQALKVRKMLAIAPEDISPVVRLVSLADSGIEENDN 68
QY 183 MYRACIATICELALQNPVVALRGGLNTILKNVIDCQLSRINEALITTLILHLNHPKTRQ 242
Db 69 LURACIATIAEFVNLNPAALLVCGVTSITRNVLECHNPRIAESLCGVLLYLLLEWFPQTRN 128
QY 243 YVRADVELERILAPYDTPFHYRHS 265
Db 129 I--CGVRLDCLAAPYCDFTYRQS 149

RESULT 4
US-09-270-767-48544
; Sequence 48544, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 48544
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48544

Query Match
  3.2%; Score 281; DB 4; Length 149;
Best Local Similarity 38.5%; Pred. No. 5e-16;
Matches 55; Conservative 35; Mismatches 51; Indels 2; Gaps 1;

QY 123 LKVDYLIARCDIQOSNEVERTQALRVKMITVNASLFPSSVTNSLAVNGDGLQERDR 182
Db 9 LQLOHLLCRSIDLMLKNDDERVQALKVRKMLAIAPEDISPVVRLVSLADSGIEENDN 68
```


Db 595 SIISKQLNDENKRRVHLEDDDDKDDITKELDVQKELIQSEVALNDL---HUTKQK--- 648
QY 489 PYSIHLPHIIQKAI-ATHQRDQYLRVQKDIIFILKO---TEEALLINLRDSQVLOHKENL 544
Db 649 ---LEDKVENLVQNLKQESNVSIQKENLELKEHIRQNEBELSRIRNE--LMQSLAQ 701
QY 545 EWNWNLGITIKPNNVNLRYKO-----EQHFRVRLLYFYKPSKLYANLDDPAKAK 599
Db 702 DSNPNFDITLKEREAVERNLKQNLSELEQNLNENKKVAF-----DVKNEN 747
QY 600 QLTUVGQFTFELSEEDGQYLEDLVKDIQVWLNASSG-MKPEPSLQNNGLITLSQH 658
Db 748 EKLVLACEDVRHLEECIAGNQLSLEKNTIVETLRKMEGEIEABLCAWKRLLEEANKY 807
QY 659 YFLFIGTSLCHPHGVKMLKSCVPQCLLNLCSLKNQDHLKLTVSSLDYSRDLGLARVILS 718
Db 808 -----EKTIEELSNARNLNTSALQLEHEHLK-----LNQKDMIEALKK 848
QY 719 KIUTAATD-----ACRLYATKHLRVLLRANVEFFNNWNGIELVLTQHDKNKTTIS--- 767
Db 849 NIEQMDTHKETQDLVSSSLEEQKLTQLINKK-----EIFIEKLERSSKLQEEL 899
QY 768 ---SEAL---DILDEACEDKANLHALQMPALSHLGDKGLLLLRP-----LSIPKG 814
Db 900 DKYSQALRKNEILLRTIEEKD--RSLGSMKEBNHLOEE--LERLREEQSRTPAPADPKT 955
QY 815 FSYLNERGYVAKOLE--KWHREYNSKVVDLIEE-----QLNEALTYTKRPVGDNYVR 865
Db 956 LDSVTTELASEVQNLITKEHLEBEIKHKQKIIDQNSQKMLQSLQEQKKEHDEFY-- 1013
QY 866 RSNQRLQRPVHYPIHLYGQLVH-HKTGCHLLEVQNIITELCRNVRTPDLDKWEIEKKLK 924
Db 1014 -QHEQMNAHTQLFLEKDEBIKSLOKT---IEQIKTQLEHERQDIQTDNSDIQETKVQS 1069
QY 925 ASLWALGNICSSWGLNLQENNVIPDILKLAQCEVLISRGTCVYVGLIATKQGCIDI 984
Db 1070 LNT-----ENGSRKHDLSKAETERLVKIGERLEIKLNEKN-----ISLTQKIDQL 1117
QY 985 LKCHNWDVAVRHKHLPVPPDDVEQLCNELSSIPSTLSLNSSESTSRHNSSESVPSSM 1044
Db 1118 SK-----DEV-----GKLTQIQKDLQIQAHLARISSTHT-QDVV 1153
QY 1045 FILEDDRFGSSSTTFDLINEDTEPTFYDRSGPIKDKNSFPFFASSKLKRNILSLTL 1104
Db 1154 YLQOQLQAVAMEREKFAVLNKTRENSH-----LKTEYHKWMDIVAAKEAALIKQD 1206
QY 1105 PNKHRSSSDPKGKLSSEKTSNRIRITLTPSVDFNHSDDTPTIPSTVQKTL--QLETS 1162
Db 1207 ENKKLSTRFESSGQDMFRE--TTQNLRIIREKDIID-----ALSQKQTLAVLQTS 1258
QY 1163 FMGNKHIEDGTSPS-----IGEND-----LKFTKNFGTENHRENTSR 1200
Db 1259 STGN-----EAGGVNSIQFPELLQERKLVKQVKMEEWKQVMTTVQNMQHSQAQLEEL 1314
QY 1201 ERL-----VVESSSTSRMKI-----RSQSFNFTDTTSGTSSMSSSPSRETGVADATMTDTC 1252
Db 1315 HQLOAQVLVDSNNKSLQVDYTGILQSYEQNETKLNFGQELAQVQHSIGQLCNTKDLLL 1374
QY 1253 GMSMTVVTYKTKTSHYLTQPSNH--LSLSKNSVSLVPPGSSHTLPPRAQSLK---APSI 1308
Db 1375 GKLDII--SPQLSSASILLTPQSAECLRAKSEVL-----ESSELQOELFELAKSLQEKD 1428
QY 1309 ATIKSLADCNFSYTSRDAFGYATLKLQOORMHPSLSHSEALAPAKOVLFTDTITMKA 1368
Db 1429 ATIRTIQENHRUSDIAA--TSELRKEHEQDTSBKQEKQDVLQKLLKEKDLIIKA 1486
QY 1369 NSPESRLTPSRFMKALSALDKEDLLSPINQNTLQRS--SSVRSVMSSATYGGSDYIG 1426
Db 1487 KS--DQLLSS-----NENFTKNVNEILLRQAVTNLKERI----- 1519
QY 1427 LALPVDIN-----DIFQVKDIPY-----FQTKNIPPHD--DRGARAF 1462

Db 1520 LILEMDIGIKLKGNEKIVETRYGKTEYQALQETNMKFMMLREKEFECHSMKEKALAFE 1579
QY 1463 HDAGGLPSGTGLVKNFSHLLRQOMSLTEIMNTHSDASLFL--ESTEDTGLQSHHTDNC 1520
Db 1580 QLLKEKEQGTGELNQLLNAVSKMQEKTVVYQOERDOVMLAKQKQMENTALQNEVOR-- 1637
QY 1521 LYCVCIIEILGFQPSNQLSAICSH-----SDFODIPYSDMCROTTHNPVLEVPSPKFGSIS 1574
Db 1638 -----LRDKFERNQBELRLNHLSEDSYTTREALAAEDREAKRKKTVLEEKLVSSS 1692
QY 1575 GCSGQVSGSASSTKSTELLGVTIPDDTPMCRILLRKEVLRLVINLSSSVSTKCHET 1634
Db 1693 NAMENASHQASQVQESLQEQNLNVVSKORDETALQLSVSOEQVQOYALSLAN----- 1743
QY 1635 GLITIKETKYPQTFDDICLYS 1654
Db 1744 -LQWVLEHFOQ--EEKAMYS 1760

RESULT 8

US-09-538-092-330
; Sequence 330, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormat Version 0.9
; SEQ ID NO 330
; LENGTH: 1306
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YGR014W
US-09-538-092-330

Query Match 2.2%; Score 196.5; DB 4; Length 1306;

Best Local Similarity 20.9%; Pred. No. 4.5e-07;
Matches 162; Conservative 119; Mismatches 270; Indels 225; Gaps 31;

QY 1004 VPDVBEQLCNELSSIPSTLSLNSSESTSRH-NSESESVSPSMFILEDDDFGSS---STST 1059
Db 157 VPID-----TSANSIPFTTTSNIETTSAPLTSDTPLISTSTWSAADNVFSSANPISASL 211
QY 1060 PFLDINEDTEPTFYDRSGPIKDKNSFPFFASSKLKRNILN----- 1100
Db 212 TTTDSSESFDQT--SPAGALPVQSSADFPSSSRILVQSSADFPSSPSTTTDLSAAPL 269
QY 1101 -----SLTLPNKHRRSSDDPKG-----GKLSSEKTSNRRI-RTLTE 1136
Db 270 QTSSESSFTTAALP-----VSSSTDVDGSSASPVVMSAAGQIASSSTDNPTMSETFSL 325
QY 1137 PSVDFNHSDFTPISTVQKTLQLETSFMGNKHIEDGTSPSIGENDLKFTKNFGTENHRE 1196
Db 326 TSTEVDGSDVSVTSALLSAPLQTS-----TSNSFSIVSPSVFVPSQSSD--- 373
QY 1197 NTSRERLVVESSSTSSHKIRQSFN---TDTTTSG-----ISSMSSSPSRETGVDA 1245
Db 374 -----VASSSTAN---VSSSFDIPQTSTSGSVVSAQSASALAFQSSSTEVYGASA 423
QY 1246 TWDTDGSGMSTVSTYKTKTSHYLTQPSNHLSLKNSVSLVPPGSSHT--LPRRAQSL 1303
Db 424 SS-----TWSSLLSTSLQST---TLDSSSLASSASSSSDLTDYGVSSSTASIPLLSASE 474

1304 KA-----PSIATIKSLADCNFSYSS-----RDAFGYATLKLQOORHPSL-----1345
475 QASTSSSFVSVSPSVFVSQSSDVASTSAPSVSSFSYSLQAGGSMTPNPSSTIV 534
1346 -----SHEALASAPAKOVLFDTTMTKANSPESR-----LTPSRF 1380
535 YSSSTGSSESAASTASATLSGSSSTYMAGNLQSQPPSTSSLLSESQATSTSAVLASSSV 594
1381 MKALSY-----ASLDKEDLLSPINQNTLQSSS-----VSMVSSATYGSD-----1422
595 STTSPYTTAGGASTEASSLISSTAETSQVSYSQSTTALQTSFASSTTEGSETSSQGF 654
1423 --DYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGGLPGSGTGVKNSF 1480
655 STSSVLVQMPSSISSEFS-----PSQTTQWNSAS 684
1481 HLLRQOMSLTEIMNSIHSASLFLSTEDTGLQEH-----CIEILGPOPSNQL 1537
685 SSSQVYTSIGILSQV-SDTSVSY-TTSSSVSVQSDTPVSYTTSSSVSVQSDTPVSYT 742
1538 SAICSHSDFDIPYSDWCEQTIHNPLEVVPKFSGICGSDGVSGSASSTKSTELLIG 1597
743 TSSSVSVQSDTPVSYTTSSSVSVQSDTPVSY-TTSSSVSVQV-SDTSPVSTSSRSS 798
1598 VKTIPDDTPMCRILLRKEVLRVINL-----SSSVSTKCHE 1633
799 VSOV-SDTPVPSTSSRSVSVQSSSLOPTTSSQRTTSTHGALESSESVSQASE 853

RESULT 9
US-09-248-796A-16699
; Sequence 16699, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16699
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16699

Query Match 2.2%; Score 196; DB 4; Length 823;
Best Local Similarity 22.8%; Pred. No. 2.2e-07;
Matches 153; Conservative 103; Mismatches 279; Indels 136; Gaps 28;

1015 LSSIPSTLSNSES-SRHNSESSEVPSSMFLDDEDRFGS--SSTSFELDINDTEPT 1071
174 VSSYDSYSSSTESSIASYSSSTIESSTLSSDRYSSISDTTSFWDSSSDLEST 233
1072 FYDRSGPIKDNKNSFPFFASSKLKVRNRLNLTPLNKHRSDDPKG-----GKLSE--- 1123
234 LITWSSSIDAQ-----SSHLVQS-VNSISLTSQELSSSSSESSSTATDALVSSDASS 285
1124 ---SKTSNRIRLTPEPSVDFNHS-----DDFTPISTVOKTQLE-----TSFMG 1165
286 ILSSTSYYPESTLTPSDDPHTTAGESDSQSISFTSVTEISSDSVSLTSDPASSFDS 345
1166 NKHI--EDTGSTPISGENDLKFTKNFGT-----ENHRENTSRERLVE 1206
346 SSSLNSDSSSLPSSQSDILLTSSSFTLVVPFSLSSSSSLTYPHYVNVSTTYHASESE 405
1207 SSTSSHKIRSQSFNTDTTTSIGISSMSSSPSRRTVGVDATTM-----DTCGSMSTVYST 1261

406 SSSVASPWSASESANDDTHT---LLESTDTTSSIGTSDSSTVTCRRDNGDCGCVTGMP 461
1262 KTI-----KTSHYL---TPQSNHLSLKSNSV-SLVPFGSSHT-----LPRRAQS 1302
462 SSIDSEQTSDDVTSSFFVASSTPTSAEQSIITDNPNDLSQTSASSSTKLVSFVSDTVVNS 521
1303 LKAPSIATIKS-----LADCNFSYSSRDAGFYATLKLQOQOORHPSLHSEALASPAKDV 1358
522 ISLSTSTLSDDSTSSDTSISSYNSDT-----GNINAGSLHTSTASIKELSTQKTGV 575
1359 LFTDTITWKANSFESRLTPSRFMKALSASLDKEDLLSPINQNTLQSSSVSRSMVSSATY 1418
576 MLSSSYLSTKLSSSTSDITTELLITTELTTELTTELTEDNEP---NTFTSTPSSHSEIFS--- 629
1419 GSGDDYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGGLPGSGTGVKNS 1478
630 --SD-----NSVLS-KQVDGESTVEIPVPTDT-TTVSSSVSVHSIEAFTATLGEN 674
1479 SFHLLRQOMSLTEIMNSIHSASLFLSTEDTGLQEH-----TDDNCLYCVCI 1526
675 SFSKVASAPVNT---TSLRSTSSSNHATESGTVKSEASAEALPSPPTSTDNRLSYSTE 732
1527 EILGQPSNQLSAICSHSDFDQDIPYSDWCEQTIHNPLEVVPKFSGICGSDGVSGQESA 1586
733 EAKGSTYANSSTNNLMTESQVAAPTDTSTSVLTANP--VVTSTFDDKS--SAAVNO---P 785
1587 SSTKSTELLIG 1597
786 SKTKSIEESIG 796

RESULT 10
US-09-054-272-38
; Sequence 38, Application US/09054272
; Patent No. 6692909
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Daley, George Q.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Rozen, Steven G.
; TITLE OF INVENTION: CODING SEQUENCE POLYMORPHISMS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/054,272
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI98-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:
; INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
LENGTH: 2224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-054-272-38

Query Match 2.1%; Score 189; DB 4; Length 2224;

Best Local Similarity 17.5%; Pred No. 5.1e-06;
Matches 301; Conservative 242; Mismatches 587; Indels 586; Gaps 81;

QY 352 SVPGR-FODSWRLS-DGFAAEAKTILPHRASRPDLMDNY----- 391
DB 141 AVAPGREYTSIESDSGPHDDPPCLTHIYSHENLIBDFNSGLGLPILCKKGLTE 200
QY 392 -----LALILSAFIR-----NGLLEGLVEVIT-NSDDHISVRATIL 426
DB 201 GGTQKTPDKQIVLLFAVDFDESKWSQSSSLMYTVNGVGTMPDITVCAHDHISWH---L 257
QY 427 LG-----EL-LHMANTILPHSHSHLHCL-----PTLNMMAA-----SFDIPKEK 465
DB 258 LHMSSGPELSIHFGQVLEQNH-HKVSATLVSATTTANMTVGPCKWIIISLTPKHL 316
QY 466 RLRSAAAL---NCLRPFHEMKK----- 490
DB 317 QAGQAVIDIKNCPPKTRNLKKTREQRHMRWEYFIAABEVIWDYAPVIPANMDKKYR 376
QY 491 SLHLDHIQKAIATHQRQDQVLRVQKDIPIFKOTEEALLNLRDSQVL-----QHKENL 544
DB 377 SOHLDN-FSNQIGHYKVMYTOYEDSF---TKHTVNPNMKEDGILGPIIRAQVDTL 431
QY 545 EWNWNLGITLK-----WPN-VNLRNFKDB-----QLHRFV 574
DB 432 K-----IVFKWASRPYIYHGVTFSPEYEVNSSFSGRNNTMIRAVQGETYTK 484
QY 575 RRLLYFKPSK-----LYANLNL--DFA-----KAKQLTVVGCQ----- 607
DB 485 WNILEFDEPTENDAQCILTRPYSDVDIMRDIASGLIGLILCKRSLSLDREGIQRAADIEQ 544
QY 608 ---FTELLSEEDGQYLEDVKDQVOMLNASGMKPSRLQNLGLTLTSHYFLFIG 664
DB 545 QAVFAVP---DENKSIDNINKFCE--NPDEVKRDDEPKFYESNIMSTINGV----- 593
QY 665 TLSCHPHGVKLEKCSVFQCLNLCSLKNQDHLLK--TVSSLDYS-----RD 710
DB 594 -----PESITTLGCFDDTVQMHFCSVGTQNEILTHFTGHSFYGRHEDTTLFPMRG 648
QY 711 GLARVILSKI---LTAATDACRLYATKHLRLVLR-----ANVEFFNNGIELLV 756
DB 649 ESVTVTMNVGTWMLTSMNSPR---SKKLRLKFRDVKCIPDDDEDSYEIFEPPESTVMA 705
QY 757 T-QLHD-----KNTISSEALDILDEACEDKANLHAL 787
DB 706 TRKHDRLEPEDESADYQNRALAAALGIRSFNSLNQE-----EEFNLTAL 756
QY 788 IQMKPALSHLGDGLLLRLFLSLPKGFSYLNERYGVAKOLEKWHR----- 833
DB 757 -ALENGTEFVSNTDIIVGNSYSPSNISKFTVNNLAEPKAPSHQOQTAGSPLRLHIG 815
QY 834 -----EYNSKY-VDLIEBQLNEALTYRKPDVG-----DNYVRSNORLORPHV 876
DB 816 KNSVLNSSTAHESSPYSEDIPLQPDVGTGIRLLSLGAGEFSPQEHAKRGKPVVERDQA 875
QY 877 YLPIHLVQVHLHKTGCHLLEVNQIITELCNVTPD-LDKWEEKLKLKASLWALNIGS.935
DB 876 AKHRFSMKLLAHKVGRR-----LSQDTGSPSGMRPEWEDLPS-----QDTGS 917
QY 936 SN-----WGLNLQEVNIPDILKLAQCEVLSIRGTCVYVGL--TAKTKQGDILKCHN 989
DB 918 PSRRPW-----EDPPSOLL-LLQSNSSKI-----LVGRWHLASEKGSYEIIQ--- 960

QY 990 WDAVRHSRKHLPVVPDDVEQLCNELSSIPSTLSLNSSESTSSRHNSSESVSPSSMFILED 1049
DB 961 -DDEDATAVNWMLISQNASRANGE-----STPLANKPGQSGHPKPRVHRHKSQVLRQD 1014
QY 1050 DRFGSSSTSTFFLDINEDTEPTFYDRSGPIKDKNSPPF--FASKLVKNRILNLTLPNK 1107
DB 1015 GGSRLKKSQFLIKTRKKKKEK-HTHAPLSPRTFHPRLRSEAYNTFSERLKHSVLV--- 1070
QY 1108 KHRSSSDPKGKGLSSKSTNRIRTLTEPSVDF-----NHSDDQFTPIST--V 1153
DB 1071 -HKS-----NETSLPTDLNQTLPDMDFGWIASLPDHNQNSNDTGOQACSPGL 1117
QY 1154 QKTLOLETSFMGNKHIEDTGSTPESIGENDLKTKNFGTENHRENTSRERLIVVESTSSHM 1213
DB 1118 YQTVPEEHYQ-FPIQDPDMHSTSD-----PSHRSSSPSELSEMLEYDRS--- 1162
QY 1214 KIRSQENTDTTTSIGISSMSSSPSRETGVGDATMTDTCGMS-----T 1257
DB 1163 ---HKSFPPTD-----ISQMSPSSEHE---YQMTVISPDLQSVTLSPQLSPDLSTHT 1211
QY 1258 VWTIKTKSHYLTPOSNHLSSKNSVSLVPPGSSHTL-----PRRAQSLKAP 1306
DB 1212 TLSPELIQRN--LSPALGQMPISPDLSHHTLSLSPDLSTHTLSLQNLSPQLSPQLSP 1269
QY 1307 SIATIKSLAD-----CNFSYTSRDAFGYATLK-----RLQQORMHPSLSH-- 1347
DB 1270 ALGQMPPLSPDLSTHTTISLQPSQNLSPQLSPQLSPQLSPQLSPQLSPQLSPQLSP 1329
QY 1348 -----SEALASPA-----KDVLFDTITMKANSFESRLTPSRFMKALS- 1385
DB 1330 LSLDFSQNLSPQLSPQLSPQLSPQLSPQLSPQLSPQLSPQLSPQLSPQLSPQLSP 1386
QY 1386 ---YASLDKEDLLSPINQNTLQRSSVRSMVSSATYGG---SDDYIGLALPVDIND 1435
DB 1387 DLSEMLFADLSQIPTPDLQMTL--SPDLGTDLSNPGQMSLSPDLQSVTLSPDLSD 1444
QY 1436 IFQVKDIPYQTKNIPPHD-----DRGARAFAH-DAGGLPFGTGGVLKN 1478
DB 1445 TTLPLDLSQIS-----PPDLQIFPSESSQSLLOEFNESFPYDGLQMPSPSSPTLND 1500
QY 1479 SFHLLQQMSLTIMNSIHSASLFLBSTDGTQLOEHTDNCILVCVIEILGFQPSNQLS 1538
DB 1501 TP--LSKEFNPLVIV-GLSKDGTDIETIIPKEEVQSSDD----- 1537
QY 1539 AICSHSDFDQIPDYDWCQETI-----HNPLBVPVSKFSGISG-----CSDGVSOEG 1584
DB 1538 ---YAEIDYVPDDPYKTDVTRTNINSSRDPDNIAAWYLRNNGRRNRYIAAEEISWDY 1593
QY 1585 SASSTKSTELLGVKTIPTDTPMCRILLRK----- 1614
DB 1594 SEFVQRETD-IEDSDDIPEDTIVKVVFRKYLDTFTKRDPRGEYEHGILGPIIRA EV 1652
QY 1615 -EVLRLVINLSSVSTKCHETGLTTIKEKYPQTFFD 1649
DB 1653 DDVIQVRFNKLASRRPYSLHAHGLSYEKSSEGKTYED 1688

RESULT 11

US-09-949-016-7876
; Sequence 7876, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: CL001307
; FILE REFERENCE: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7876
;; LENGTH: 3854
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-7876

Query Match 2.0%; Score 179.5; DB 4; Length 3854;
Best Local Similarity 17.8%; Pred. No. 8.9e-05;
Matches 339; Conservative 245; Mismatches 612; Indels 713; Gaps 83;

QY 45 VARLOQVSNRKLGNLNNFTKLLCDIGHSEKLGPHVEDIIC---LRLALNEAKEVRA 101
DB 1824 VCFPMKITVIGRYGSTNARAKI-----PLGFYGYHTYILPWESEKLMHD--PLKG 1872

QY 102 AGLRALRYLIQSSILQKVLKVDYLIARCIIDIOQSNVEVTOALRLVRKMITVNASLF 161
DB 1873 EG-----ESANQPEIDQHLAMVVALQEDIQCRYNLACHRLLETLLQ-SIDL 1917

QY 162 PSSVTNS---LIAVNGDGLQERDMVRA---CIAICELALONPEVVALRGGINILKVV 215
DB 1918 PLSNANAOYFLRKKPKAVEEDSRVESAYODCICQLQLNLAHNAVORLKVAGASRKM 1977

QY 216 IDC-----QLSRINEALITITLILHNLHPKTRQYVRADVLELERILAPVTDPHY 262
DB 1978 SETSNPEDLIQTSSTQELATIRYLLDTLLSL-HASNGHSPVAVLQ----- 2023

QY 263 RHSPDTAEGOLKEDREARFLASKMGIIATRSWAGIINLCKPGNSIGIQLGVLCI--P 319
DB 2024 -----STFHAQA-----CEELPKHLICISGTP 2044

QY 320 NMEIRGLL-----EVLVDIIFRL-PLPVVTEFEALISV-----DP 355
DB 2045 KIRLHTGLLLVLQCGGERWGFSLNVLOELYNSEQLLIFPDQDRVPMLLSCIGQRSLSNS 2104

QY 356 GRFQDSWRLSDGFVAEAKTILPHRARSPLD---MDNYLALILSAFIRNGLEGLVEVI 412
DB 2105 GVLESILNLLDNL-SPLOQLPMHRTGVLDPIMISVMVLSR-----LLDYV 2154

QY 413 TNSDDHISVRAATILGELLHMANTILPHSHSHLHCLPTLMNMAASFDIPKEKRLASNA 472
DB 2155 ATVEDEAAAAKPLNGNQWSFINNLL---HTQSLN-----RSSKG 2191

QY 473 LNCLEKRFHEMKRGKPKYSLDHLIIQKAIATHQKEDQYLRVQKIDIFILKOTEEALLNL 532
DB 2192 SSSLRLYSRK-----IRKOLVHHKQO-----LNLKAKQKALVBOQ 2228

QY 533 RDSQVLQHKENLEWNNWNLITGLKWPNNVNLRYKDEQLHRFVRRLLYFYKPSKLYANLD 592
DB 2229 EXEKIQSNKS---SYKLLAVEQAKLQATSKHFKD-----LIRLRRTAEMERSNLD 2276

QY 593 LDFAKAKQTVVGCQPTFELLESDEGGYLEDLVKDIVQWLNASSGMKPKERSLQNGLL 652
DB 2277 TEVTAK-----ESPE-----IBPLPPTLAH-----ERCI----- 2301

QY 653 TTLSOHYFLFICTLSCPHGVKMLEKCSVFQCLLN-----LCSIKNODHLKLTVSSL 705
DB 2302 -SVQKLVLLLSMDPTCHADLLLVCKVLARIANATPTIHLCEIVNEPQLERLL----- 2356

QY 706 DYSRDGLARVILSKIIATADCLRYATKHLRVLLRANVEFFNNNGIELLVTLQHDKNKT 765
DB 2357 -----LLLVGTD-----FNRGDISWGGAAQYSLTCLQD---I 2387

QY 766 ISSEAL-DILDEACEDKANLHALIQMKPALSHLGDKLLLLRLFLSIPKGFSLYNERGV 824
DB 2388 LAGELLAPVAEAMEE-----GTVGD-----DVGAT 2413

QY 825 AKQLEKWHREYNSKYVDLIEEQLEALITTYRKPVDGDNVVRNSQRLQRPVHYLHYG 884
DB 2414 A-----GD-----SDSLQQSSVQLLETIDE 2434

RESULT 12

US-07-741-940-7

; Sequence 7, Application US/07741940

; Patent No. 5352775

GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JESLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,940
FILING DATE: 19920109
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.035574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
US-07-741-940-7

Query Match 2.0%; Score 176.5; DB 1; Length 2842;
Best Local Similarity 17.9%; Pred. No. 9.4e-05;
Matches 360; Conservative 266; Mismatches 644; Indels 737; Gaps 96;
QY 58 GHLNFTK---LLCDIGHSEKLGIFYEDI-----IICLRALINE 95
DB 132 GYLELEKERSULLADLDKEEKDYQAOLNLTKRIDSLLTENFSLQDTRQLEYE 191
QY 96 AKEVRAA-----GLRALYLIQDSSILQKVLKVDYLIARCI----- 133
DB 192 ARQIRVAMEQLGTCQDMKEAQRRIARIQIEKIDILRIIR-QLLSQQATEAERSQNKHE 250
QY 134 ---DIQOSNEVERTQALRL-----VRKMTVNASLFPSSVTN----- 167
DB 251 TGSHAERQNEGGVGGEINMATSNGQSGTTRMDHETASVLSSTTSAPRRLTSHLGTK 310
QY 168 ---SLIAGVNDGLQERDMVR-----ACIAICELALONPEVVAL 204
DB 311 VEMYVSLISM--LGTHDKDDMSRTLLAMSSQDSCISNRQSGCCLPLIQLLHGNDKDSVL 368
QY 205 RGG-----LNTILKNVIDCQLGRINELATITLHLNHPKTRQYVRADVELERIL 254
DB 369 LGNSRGSKEARASAAALHNIHSQPDOKRGRREIRVLHLE--QIRAYCETCWQEAH 426

QY 255 APYTDFHYRHSPTAEQ-----LKEDREARFLASKMGII----- 289
DB 427 EPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEBHRHAMNELGGLQAIALLQVDCEMYG 486
QY 290 -----ATFRSWAGI-----INLC-----KPGNSIGSLGLVL 316
DB 487 LTNDHYSITLRRYAGMAL/TNLTFGDVANKATLCSMKCMGRALVAQLKASESEDLOQVITASV 546
QY 317 CIPNMEIR-----RGLLEV--LYDIFRLPLPVVTEEFIBALLSVDPGRFQDSWRLS--- 365
DB 547 -LNLNLSRADVNSKKTIREVGSVKALMECALEVVKESTLKSLSA-----LNLNLSAHC 598
QY 366 -----DGFVAEAKTILPHRARSFPLDMNDNYLILSAFIRNGLLEGLVEYITNS 415
DB 599 TENKADICAVDGAFLVGT-LTVRSOT-----NTLAIIESG--GGILNRVSSLIATN 648
QY 416 DDHISVRATILLGELLHMAN--TILPHSHSHL-----HCLPTLMNMAASFDIPKXR-L 467
DB 649 EDH-----RQILFENNCLQTLLOHLKSHSLTIVSNACGTLNLSAR--NPKDQEAL 697
QY 468 RASAAALNCLKRF---HEMKRG-----PKPY-----SLHLDHII 498
DB 698 WDMGAVSMKXNLHSHKMTAMGSAALRNLNMANRPKYKDANIMSPGSSLPSSHVRK-- 755
QY 499 QKAI-----ATHQKRDQYLRVQKDIIFILKDTTEBAILINLRDSQ 536
DB 756 QKALEAELDAQHLSETFDNIDNLSPKASHRSKQHQSLGYDYP-DTN----- 803
QY 537 VLOHKENLEWNWNLIGTILKWPNNL-----RNYKDEQLHRFVRRLLYF 580
DB 804 --RHDDNRSDNFNTGNMTVLSPYLNTVLPSSSSSGSLDSSRSEKDRSLRERERIGLGN 861
QY 581 YKPSSKLYANLDDFAKAKQLTVVGCOPTFLE-----SEEDGQVLEDLVKDIQVWL 634
DB 862 YHPATE---NPGTSSKRGQLISTTAAQIAKMEEVSAIHTSQED----- 902
QY 635 NASSGMKPE-----RSIQNQLLTLSQHYFLFGLTSLCHPGVKMLEKCSVFCOLL 686
DB 903 -RSSGSTTELHCVTDERNALRRSSAAHTNTY----- 934
QY 687 NLCSLKNQDHLKLTVSSLDYSR---DGLARVILSKILTATACRLYATYKHLRVLLRAN 743
DB 935 NFTKSENSNRTCSMPYAKLEYKSSNDSLSNV-----SSDG---YKG---RGQMKPS 981
QY 744 VEPFNWGIELVTLQHDKNKTISSERALDILDEACEDKAN-LHALIQMKALSHLQKGL 802
DB 982 IESYSE-----DDESKFCS-----YGYPADLAHKIH-----SANHMDND- 1017
QY 803 LLLRLFLSI PKGFS-----YLNERGVAQOLEKWHREYNSKYVDLIEEOLNEA----- 850
DB 1018 -----GELDTPIVLSKYSDQLNSGRQSPSONERWAPKH-----IIEDIKGEQQR 1068
QY 851 --LTTYRKPDGDNVRRSNRQLRPHVYLIHYGQ-----LVHHKTGCHLLEV----- 898
DB 1069 NQSTTY--PV---YTESTDDK---HLKFQPH-FGQCECVSPVRSRGANGSETNRVGSN 1117
QY 899 ---QNIITELCRNVRTPDLDKWEETKKLKASLWALGNIGSSNWGLNLLQENVIPDLK 954
DB 1118 HGINQNVQSLOCQ-----DDYEDDKPTNY-----SERYSEEQHEEERPTNYS 1162
QY 955 LAQCEVLSIRGTCCVVLGLIAKTQGCODILKCHNWDVHRSHKMLWPVVDVDEQLCNE 1014
DB 1163 IKNEE-----KKHVQPIDISLUKIATD-IPSSQKQ----- 1192
QY 1015 LSSIPSTLSINSESTSRH---NSESESVPSSMFILEDDRFSGSSTSTF-----F 1061
DB 1193 --SFSFSKSSGSGOSTEHSMSSESTSTPSSNAKRNQNLHPSSAQSRSPQPKAATCKV 1250
QY 1062 LDINEDTEPTFYDRSGPIKDKNSFPFFASSKLVKNRILNLSLTLPNKHRSSSDPKG-GKL 1120
DB 1251 SSINQETIQTVCYVEDTPI---CFSRCS-----LSSLS-----SAEIDEIGNQT 1291

1121 SSESKTNRIRLTPEPSVDFNHSDDFTPISTVQKTLQLE--TSFMGNKHIED-TGSTPS 1177
1292 TQADGAN-----TLQAEIKKIGTSAEDPVSEVA 1324
1178 IGEND-LKTYKNGTENHRENTSREBLVESSTSHMKIRSQS-----FNTDTTTS 1227
1325 VSQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSQAOTPKSPPEHYVQETPLMFS 1384
1228 GTSSMSSPSRETVGVDATMTDTC--GSMSTVVSTTKTSHVLTTPQSNHLSLSKSNVS 1286
1385 RCTSVSSLSFERSSTASSVQSEPCGWSGISPSDL-----PDSGQTPMPSRSKT 1437
1287 LVPPGSHLTLPRAQSLKAPSIATIKSLADCNFSYTSRDAFYATLKRLOQOORMPSPS 1346
1438 PPPPQTAQTKREVKNKAPTAEKRES-----GPKQAAVNAVQRVQ----- 1479
1347 HSEALASPAKDVLTFTTITWANKANSFESRLTPSRP--MKALSYASLDKE-----DLIS 1396
1480 -----VLPDADTLL-----HFATESTPDGFCSSLSALSLEDPFIQKDVLRIMP 1525
1397 PINON-----TLQRSSVRSMVSATYGGSDDYIGLALPVDINDIFQVKDIPYFQT 1447
1526 PVOENDNGNETSEQPKENENQEAETIIOSEKDLL-----DOSDDDDIIEIEECII 1579
1448 KNIPPHDDRGARAFADAGLPSGTGGLVKNSPHLLRQOMSLTEIMNSIHSASLPLEST 1507
1580 SAMPTKSRKAKKPAQATASKLPPVAR--KPS-----QLPVYKLLPSONR-----LQPK 1626
1508 EDTGLEHTDDNCLYCVCHIEILQFQPSNOLSAICSHSDFDQDIPYSWCQCTHNPLEVVP 1567
1627 KHVSFTPGDMPRVYCVBGTPIFNSTATSL-----DLTIESP-----P 1665
1568 SKFSGISGCGDVGVSQES-----ASSTKSTELLGKVTIPDDTPMCRILLRKEVLRIV 1620
1666 NEIAAGEVRGG-AQSGEFKEKRTIIEGRSTDEAOGGKT----- 1704
1621 INLSSSVS-----TKCHETGLLTIKKYPOTPDIDICLYSEVSHLSHCTFLPCRRFIQ 1674
1705 -----SSVTIPELDNKAEGDILA-----ECINSAMPKGSHPFR-----VK 1743
1675 ELFDQVQFLOMHEEAVALTPPKQPI 1701
1744 KIMDQVQ-----QASASSAPKNQOL 1764

RESULT 13

US-08-289-548A-7
Sequence 7, Application US/08289548A
Patent No. 5648212

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, JOHNA
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESS: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,548A
FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.46943
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
US-08-289-548A-7
Query Match 2.0%; Score 176.5; DB 1; Length 2842;
Best Local Similarity 17.9%; Pred. No. 9.4e-05;
Matches 360; Conservative 266; Mismatches 644; Indels 737; Gaps 96;
QY 58 GHLNNFTK---LLCDIGHSEKLGFIHYEDI-----II CLRALLNE 95
DB 132 GYLEELKERSLLADLDKEEKDQVYAOQLNLTNRIDSLTENFSLQDTMTQRLVE 191
QY 96 AKEVRAA-----GLRALVLIQDSSILQVLLKLVLYLARC----- 133
DB 192 ARQIRVAMEEQLGTQCDMEKRAQRIARIQIQEKDILRIR-QLLOSQATEAERSQNKHE 250
QY 134 ---DIQOSNEVERTQALRL-----VRKMTVNASLPPSVTN----- 167
DB 251 TGSHTDAERQNEGGVGEINMATSNGQGGSTTRMDHETASVLSSTHAPRLTSHLGTK 310
QY 168 ---SLIAVGNDGLQERDMVR-----ACIAIICELALQNPVEVAL 204
DB 311 VEMVYSLLSM--LGTHDKDDMSRTLLAMSSQSDSCISMESQSCLPILLIQLLHGNDKDSVL 368
QY 205 RCG-----LNTILKNVIDCQLSRINEALITILHLLNHPKTRQVVRADVELERIL 254
DB 369 LGNSRGSKEARASAAALHNIHSQPDCKRGRREIRVLHLE--QIRAYCETCWEQEAH 426
QY 255 APYTDPHYRHSPTAEGO-----LKEDREARFLASKMGII----- 289
DB 427 EPGMDQDNMPAPVEHOICPAVCVLMKLSFDEHRHMANELGGLQATAELLQVDCMYG 486
QY 290 -----ATFRSWAGI-----INLC-----KFGNSGIQSLIGVL 316
DB 487 LTNDHYSITLREYAGWALTNLTFDGVANKATILCSMKGCMRALVAQLKSESEDLQQVIASV 546
QY 317 CIPNMEIR-----RGLLEV--LYDIFRLPLPVVTEBFEALLSVDPGRFQDSWRLS---- 365
DB 547 -LRNLSWRADVNSKKTLEVGSVKALMECALEVKESTLKSVLSA-----LWNLSAHC 598
QY 366 -----DGFVAEAKTILPHRARSRLMDNYLALILSAFIRNGLLEGLEVITVS 415
DB 599 TENKADICAVDGAFLVGT-LTYRSQT-----NTLAIIESG---GGILRVNSSLIATN 648
QY 416 DDHISVRATILGELLHMAN---TILPHSHSHL-----HCLPTLMNMAASFDIPKEKR-L 467
DB 649 EDH-----RQILRENNCIQTLLQHLKSHSLTIVSNACGTLWNLSAR--NPXQDEAL 697
QY 468 RASAALNCLKRF-----HEWKKRG-----PKPY-----SLHLDHII 498
DB 698 WDMGAVSMLKNLHSHKHMIAMGSAALRNLMANRPKDYKDNIMSPGSSLSPLSHVRK-- 755


```
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
US-08-452-654-7

Query Match      2.0%; Score 176.5; DB 1; Length 2842;
Best Local Similarity 17.9%; Pred. No. 9.4e-05;
Matches 360; Conservative 266; Mismatches 644; Indels 737; Gaps 96;

QY 58 CHLNNFTK---LLCDIGHSEKLGPHYEDI-----IICLRALINE 95
DQ 132 GYLEELKERSLLADLKEKEKOWYAQLQNLTKRIDLSTENFSLQDMTRRLQLEYE 191
QY 96 AKEVRAA-----GLRALRYLIODSSILQKVLKVDVLIARCI-----133
DQ 192 ARQIRVAMEQLGTCQDMKRAQRRIARIQTEKIDILRIR-QLLQSOATEAERSSONKHE 250
QY 134 ----DIQOSNEVERTQALRL-----VRKMITYNASLFPSSVTN-----167
DQ 251 TGSHTDAERQNEGQVGEINMATSNGGQSTTRMDHETASVLSSSTHSAAPRLTSHLGTK 310
QY 168 ----SLIAGNDGLOERDMWR-----ACIALICELALONPEVAL 204
DQ 311 VEMVYLLSM--LGTHDKODMSRTLLAMSSODSCISMROSGCLPLLIQLLHGNKDSVL 368
QY 205 RGG-----LNTILKNVIDCOLSRINEALITLILHLNHPKTRQVVRADVELERIL 254
DQ 369 LQNSRGSKEARAPASAAALNHIHSQPDQKRGREIRVLHLE--QIRAYCETCWEQEAH 426
QY 255 APYTDPHYRHSPTABGQ-----LKEDREARFLASKMGII-----289
DQ 427 EPGMDQKNMPAPVPEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIELALQVDCMYG 486
QY 290 ----ATERSWAGI-----INLC-----KPNSGIQSLIGVL 316
DQ 487 LTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSESEDLQQVIASV 546
QY 317 CIPNMEIR-----RGLLEV--LYDIFRLPLPVVTEEFTEALLSDVDPGRFQSWRLS---365
DQ 547 -LRLNSWRADVNSKTLREVGSVKALMECALEVKKESTIKSVLSA-----LWNLSAHC 598
QY 366 -----DGFVAEAKTILPHRARSPDLMDNYALILSAFIRNGLLEGLVEVITNS 415
DQ 599 TENKADICAVDGAFLVGT-LTYRSQT-----NTLAIIESG-----GGILRVSSLIATN 648
QY 416 DHIISVRATILGELLHMAN--TILPHSHSHL-----HCLPTLMMMAASFDIPKEKR-I 467
DQ 649 EDH-----RQILRENNCLQTLQHLKSHSLTIVSNACGTLWNLARS--NPKDQDEAL 697
QY 468 RASAAALNCLKRF---HEMKRG-----PKPY-----SLHLDHII 498
DQ 698 WDMGAVSMLNLIHSHKMIAMGSAALNLMANRPKYKDANIMSPGSSLSLHVRK---755
QY 499 QKAI-----ATHQKRDQYLRVQKDFILKDTBEALLINLRDSQ 536
DQ 756 QKALEAELDAQHLSETFDNIDNLSPKASHRSQRHKQSLYGDYVF-DTN-----803
QY 537 VLOHKNLEWNNLIGITILKWNVN-----RNYKDBQLHRFVRRLLYF 580
DQ 804 --RHDDNRSDNFNTGNMTVLSPLNTVLPSSSSSRGSLDSRSEKORSERERIGLGN 861
QY 581 YKPSSKLYANLIDLDAKAKQLTVVGCQFTEFILLE-----SEEDGGQLEDLVKDIQVWL 634
DQ 862 YHPATE--NPGTSKRGLQISTTAAQIAKMWEEVSAIHTSQED-----902
QY 635 NASSGMKPE-----RSLQNNGLLTLSQHYFLFIGTSLCHPHGVKMLKCSVFQCLL 686
DQ 903 -RSSGTTTELHCVTDERNALRRSSAAHTSNTY-----934
QY 687 NLCSLKNQDHLKLTVSSLDYSR---DGLARVILSKILTAATDACLKLYATKHLRVLLRAN 743
DQ 935 NFKTSENRTCSMPYAKLEYKRSSNDLSNV-----SSDG---YKG---RGQMKPS 981
```

```
744 VEFFNNWGIELIVTQHDKNKNTISSEALDILDEACEDKAN-LHALIQMKPALSHLGKGL 802
982 IESYSE-----DDESKFCS-----YQYVPADLAHKIH-----SANHMDND- 1017
803 LLLLFLSIPKQFS-----YLNERYGVAQLEKWHREYNSKYVDLIEEQLNEA-----850
1018 ---GELDTPINYSLSKYDQELNSGRQSPQNERWARPKH-----IIEDEIKQSEORQR 1068
851 --LTYRKPVDGDNVRSRQRLQRPVYLPPLHLYGO-----LVHHKTCGCHLEV-----898
1069 NOSTTY--PV-----YTESDDK---HLKFQPH-FGQECVSPYRSRGANGSETNRVGSN 1117
899 ---QNIITELCRNVRTPDLDKWEIKKILKASLWALGNIGSSNWGLNLQIENVPDILK 954
1118 HGINQVQSCLQE-----DDYEDDKPTNY-----SERYSEEQHEEERTNYS 1162
955 LAKQCEVLSIRGTCVYVGLIAKTQGGDILKCHNWDVAVRHSRKHLMWPVDPDQOLCNE 1014
1163 IKYNEE-----KRHDVQPIDYSLKYATD-IPSSQKQ-----1192
1015 LSSIPSTLSNSESSTSRH---NSESESVPSMFLEDDRFSGSSSTSF-----F 1061
1193 --SFSKSSSSQSGSKTEHMSSESSENTTPSSNAKRQNLHPSSAQSRSGQPKAATCKV 1250
1062 LDINEDTEPTFYDRGPIKDKNSPFFASSKLKVNRIILNLSLTPNKKHRSSSDPKG-GKL 1120
1251 SSINQETIQTVCVETPI---CFSRCSS-----LSSLS-----SAEDEIGCNOT 1291
1121 SSESKTNRIRINTLTPSVDFNHSDDFTPISTVQTKLOLE--TSPMGKHIED-TGSTPS 1177
1292 TQEADSAN-----TLQIAETIKKIGTRSAEDPVSEVPA 1324
1178 IGEND-LKFTKNFGTENHRENTSRRLVVESTSTSHMKIRSOS-----FNTDTTTS 1227
1325 VSQHPRTKSSRLQSGSSLSSESARHKAVFSSGAKSPSGAGTPKSPPEHYVQETPLMES 1384
1228 GTSSMSSSPSRETGVGDATMTDTC-GSMSTVTVSTKTIKTSHTYLPQSNHLSLSKSNYS 1286
1385 RCTSVSSLDSFESRIASSVQSEPCSGMVGSIISPSDL-----PDSFGQTMPFSRST 1437
1287 LVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSTSSRDAPGYATLKRLOQRMHPSLS 1346
1438 PPPPQTAQTKEVPKNAKPTAEKRES-----GPKQAAVNAVQVRVQ-----1479
1347 HSEALASPAKQVLFDTTITMKANSFESRLTPSRF--MKALSVASLDKE-----DLLS 1396
1480 -----VLPDADTLL-----HPATESTDPGFCSSSLSALSDEPFIQKDVLRIMP 1525
1397 PINQN-----TLQRSSSVRSVMVSATYGGSDDYIGLALPVDINDIFQVKDIPYFOT 1447
1526 PVQENDNGNETESEQPKESNENQKEAETIDSEKDLL-----DSDDDDDIIELEECII 1579
1448 KNIPPHDRGARAFADAGGLPSGTGGLVKNKSFHLLRQOMSLTEIMNSIHSASLPLEST 1507
1580 SAMPTKSRKAKKPAQTASKLPPPVAR--KPS-----QLPVYKLLPSQNR-----LOPQ 1626
1508 EDTGLQEHDDNCLYCVCIILGFQPSNQLSAICSHSDQDIPYSDWCQETIHNPLEVVP 1567
1627 KHVSFTPGDMERVYCVGTPINFSTATSLS-----DLTIESP-----P 1665
1568 SKFSGISGSCDVSQEGS-----ASSTKSTELLIGVKTIIPDDTPMCRILLARKEVLRIV 1620
1666 NELAAGEVRGG-AQSQGEFEKEDTIPTEGRSTDEAQQGT-----1704
1621 INLSSSVS-----TYKHETGLLTTIKKYPQTFDDICLYSEVSHLLSHCTFLPCRFRFQ 1674
1705 ----SSVTIPELDDNKABEGDILA-----ECINSAMPKGSKHKPPR-----VK 1743
1675 ELFQDVQVFLQHEEAELVATPPKQPI 1701
1744 KIMDOVQ-----QASASSAPKNQOL 1764
```



```

QY 1015 LSSIPSTLSNBESTSRH---NSESEVPSMFILEDDRFSSSTSTF-----P 1061
Db 1193 --SPFSKSSGSSGSSSTHSSSENTSTPSSNAKQNLHPSSAGSQGPQKAATCKV 1250
QY 1062 LINEDTEFTFYDRGPIKDKNSFFPFASSKLVKNRILNSLTLPNKKHRSDDPKG-GKL 1120
Db 1251 SSINOETIQYCVEDTPI---CFRCS--LSSLS-----SADEIGCNOT 1291
QY 1121 SSESKTNRIRLTPEPSDFNHSDFTPISTVQKTLQLE--TSFMGNKHIED-TGSTPS 1177
Db 1292 TQEAUSAN-----TLQIAEIKKIGTKSAEDPVEVPA 1324
QY 1178 IGEN-D-LKFTKNFTEHRENTSRRLVVESSSTSHMKIRSOS-----FNTDTTTS 1227
Db 1325 VQHPRTKSRLLQSSLSSESARHKAHAFSSGAKSPSKGAQTPKSPPEHYVQETPLMES 1384
QY 1228 GISSMSSPSRETGVDAATMDTDC-GSMSTVVSTKTKTSHYLTPOSNHLSLSKSNVS 1286
Db 1385 RCTSVSLDSFESRSIASSVQSEPCSGMVGIISPSDL-----PDSFGOTMPPSRSKT 1437
QY 1287 LVPPGSSHTLPRAQSLKAPSIATIKSLADCNFSYTSRDARFYATLKLQOQRMHPSLS 1346
Db 1438 PPPPQATQKEVPKNKAPTAKRES-----GPKQAAVNAVQRVQ----- 1479
QY 1347 HSEALASPAKDLFTDTITMKANSFESRLTPSRF--MKALSYASLDKE-----DLIS 1396
Db 1480 -----VLPDADTLL-----HFATESTPDGFCSSLSALSLEPPFIQKDVLRIMP 1525
QY 1397 PINQN-----TLQSSSVRSVMSSATYGGDDYIGLALPVDINDIFQVKDIPYFQT 1447
Db 1526 PVQENDNGNETEQPKESNENQEKAEKTIIDSEKDLL-----DSDDDDIIELECII 1579
QY 1448 KNIPPHDDRAGARAFADAGLPGSGTGLVKNPHLLRQOMSLTEIMNSTHSDASLFLEST 1507
Db 1580 SAMPTKSSRKAKKPAQTASKLPPVAV--KPS-----QLPVYKLLPSQNR-----LOPQ 1626
QY 1508 EDTGLQEHDDNCLYCVCIIEILGFQPSNOLSAICSHSDFDQIPYSDWCEQTIHNPLEVVP 1567
Db 1627 KHVSFTPGDDMPRVYCVETPINFSTATSL-----DLTIESP-----P 1665
QY 1568 SKPSGISGCGSDGVSGS-----ASSTKSTELLGVKTIPTDDTFCMCRILLRKEVLRIV 1620
Db 1666 NELAAAGEVVRGG-AQSGEPEKRTIPTEGRSTDEAQQGKT----- 1704
QY 1621 INLSSVS-----TKCHETGLTIKYPQTFDDICLYSEVSHLLSHCTFELPCRRIQ 1674
Db 1705 -----SVTIPELDDNKAEEDILA-----ECINSAMPKKGKSHKPPR-----VK 1743
QY 1675 ELFDQVQFLQMHHEAEAVLATPPKOP1 1701
Db 1744 KIMDQVQ-----QASASSAPKNQOL 1764

```

Search completed: April 9, 2005, 15:16:05
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2005, 19:02:52 ; Search time 90 Seconds
(without alignments)
7339.853 Million cell updates/sec

Title: US-10-782-244-3

Perfect score: 8811

Sequence: 1 MAATIGRSLKLNLRVRGRND.....AAEVLATPPKQPIVDTSAES 1708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8811	100.0	1708	8	ADR44436
2	6632	75.3	1275	8	ADQ89852
3	4014	45.6	791	7	ADQ89852
4	1714	19.5	336	7	ADC31052
5	1661.5	18.9	1936	8	ADQ89858
6	1660	18.8	1922	4	ABB63631
7	1631	18.5	329	7	ADC49296
8	1535	17.4	313	8	ADR10059
9	1127	12.8	394	4	ASG01058
10	1127	12.8	394	4	ASG16310
11	1127	12.8	394	7	ADC32769
12	720.5	8.2	1430	6	ABR53386
13	720.5	8.2	1430	7	ADK63324
14	692	7.9	141	3	AAU37120
15	634.5	7.2	200	4	ABG01059
16	615	7.0	123	3	ABG00938
17	517	5.9	102	8	ADR09113
18	281	3.2	56	8	ABO59171
19	225	2.6	2297	7	ABW01174
20	220.5	2.5	2344	4	AAU37120
21	217	2.5	2271	6	ABU16000
22	214	2.4	2261	6	ABU18914
23	212	2.4	2137	5	ABP39618
24	212	2.4	2137	8	ADS05656
25	212	2.4	2319	7	ADJ70269

ALIGNMENTS

RESULT 1

ADR44436
ID ADR44436 standard; protein; 1708 AA.

XX AC ADR44436;

XX 18-NOV-2004 (first entry)

DT Human p200 protein SEQ ID NO:3.

DE Human p200 protein SEQ ID NO:3.

XX mTOR-associated protein; mTOR-AP; cytosolic; antidiabetic; gene therapy;

KW mTOR-AP modulation; cancer; tumour; diabetes; human; p200.

XX Homo sapiens.

XX WO2004074448-A2.

XX 02-SEP-2004.

XX 18-FEB-2004; 2004WO-US004821.

XX 18-FEB-2003; 2003US-0448035P.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Sabatini DM, Kim D, Sarbassov DD;

XX WPI; 2004-635565/61.

XX N-PSDB; ADR44434, ADR44435.

XX New mTOR-associated protein (mTOR-AP) polypeptide, useful for treating or

XX preventing a disorder that is responsive to mTOR-AP modulation, e.g.

XX cancer or diabetes.

XX Claim 1; SEQ ID NO 3; 91pp; English.

XX The present invention describes an isolated mTOR-associated protein (mTOR

XX -AP) polypeptide comprising a sequence that is 70% identical to the 1708

XX amino acid sequence of SEQ ID NO:3 (ADR44436, P1) or 70% identical to the

XX 327 amino acid sequence of SEQ ID NO:6 (ADR44439, P2), or that is encoded

XX by a nucleic acid that hybridises under high stringency conditions to a

XX complement sequence of the 951271 nucleic acid sequence of SEQ ID NO:2

XX (ADR44435, S1) or the 981 nucleic acid sequence of SEQ ID NO:5 (ADR44438,

XX S2). Also described: (1) an isolated antibody, or its fragment, which is

XX specifically immunoreactive with an epitope of an amino acid sequence of

XX P1 or P2; (2) the isolated nucleic acid encoding the polypeptide, or its

XX complement; (3) a vector comprising a nucleic acid sequence encoding an

XX mTOR-AP polypeptide; (4) an isolated host cell comprising the nucleic

Adq97333 Human can
Ada89591 Staphyloc
Abm72734 Staphyloc
Abp56876 Staphyloc
Adb75595 Prostata
Abu43109 Protein e
Abj19019 Pathogen
Adg97330 Mouse can
Adg56031 Rat Prote
Adg56035 Rat Prote
Adn23131 Bacteria
Abw01173 Candida a
Ado28932 Mouse nov
Aau34408 Staphyloc
Aau34408 Staphyloc
Abu42529 Protein e
Abg06301 Novel hum
Adp54514 Human pro
Adp23752 PRO polyp
Aaw04254 Human Pac

acid; (5) a method of producing an mTOR-AP polypeptide; (6) a method for detecting the presence of an mTOR-AP polypeptide; (7) a kit for detecting an mTOR-AP polypeptide comprising the antibody and a detectable label for detecting the antibody; (8) a method for detecting the presence of the nucleic acid in a sample; (9) a kit comprising the probe or primer nucleic acid and instructions for use; (10) an isolated, purified or recombinant complex comprising an mTOR polypeptide and an mTOR-AP; (11) a method for identifying a compound which modulates activity of an mTOR-AP polypeptide; (12) a method of inhibiting aberrant activity of an mTOR-AP-expressing cell; (13) a method of treating or preventing a disorder that is responsive to mTOR-AP modulation, in a subject; and (14) a transgenic mouse having germline and somatic cells comprising a chromosomally incorporated transgene that disrupts the genomic mTOR-AP gene and inhibits expression of the gene, where the disruption comprises insertion of a selectable marker sequence resulting in the transgenic mouse exhibiting increased susceptibility to the formation of tumours as compared to the wild type mouse. mTOR-AP sequences cytostatic and antidiabetic activities, and can be used in gene therapy. The polypeptides, polynucleotides, compounds and methods are useful for treating or preventing a disorder that is responsive to mTOR-AP modulation, e.g. cancer or diabetes. The compound, which modulates an mTOR-AP activity or expression is useful for the manufacture of a medicament for treating a disease affected by aberrant mTOR-AP activity or expression. The present sequence represents human p200, which is an mTOR-AP used in the exemplification of the present invention.

XX Sequence 1708 AA;

Query Match 100.0%; Score 8811; DB 8; Length 1708;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAIGRSLKRLVRGRNDSGEENYPLDTREPSNLRLEILQNVARLOQVSNMRKLGHL 60
 DB 1 MAAIGRSLKRLVRGRNDSGEENYPLDTREPSNLRLEILQNVARLOQVSNMRKLGHL 60

QY 61 NNFTKLLCDIGHSEELGPHYEDIIICLRALINEAKEVRAAGLRALRYLIQDSSILQKV 120
 DB 61 NNFTKLLCDIGHSEELGPHYEDIIICLRALINEAKEVRAAGLRALRYLIQDSSILQKV 120

QY 121 LKLVLYLIARCIIDIOQSENEVERTQALRLVRKMITVNASLFPSSVTNLSIANGDGLQER 180
 DB 121 LKLVLYLIARCIIDIOQSENEVERTQALRLVRKMITVNASLFPSSVTNLSIANGDGLQER 180

QY 181 DRMVRACIALICELAQNPVEVALRGGLNTILKNVIDCOLSRINEALITILHLNHPKT 240
 DB 181 DRMVRACIALICELAQNPVEVALRGGLNTILKNVIDCOLSRINEALITILHLNHPKT 240

QY 241 ROYVRADVELERILAPYTFHYRHSPTAEGQLKEDREARFLASKMGIIATFRSWAGIIN 300
 DB 241 ROYVRADVELERILAPYTFHYRHSPTAEGQLKEDREARFLASKMGIIATFRSWAGIIN 300

QY 301 LCKPGNSGIQSLIGVLCIPNMEIRRGLEVLVDIFRLPLPVVTEEFIEALLSVDPGRFQD 360
 DB 301 LCKPGNSGIQSLIGVLCIPNMEIRRGLEVLVDIFRLPLPVVTEEFIEALLSVDPGRFQD 360

QY 361 SWRLSDGFVAAEAKTILPHRARSRPLMDNYLALISAFIRNGLLEGLVEVITNSDDHLS 420
 DB 361 SWRLSDGFVAAEAKTILPHRARSRPLMDNYLALISAFIRNGLLEGLVEVITNSDDHLS 420

QY 421 VRATILIGELLMANITILPHSHSHHLCPLTMNMAASFDIPKEKRLASAALNCLKRFH 480
 DB 421 VRATILIGELLMANITILPHSHSHHLCPLTMNMAASFDIPKEKRLASAALNCLKRFH 480

QY 481 EMKRGPKPYSYLDHIIQKAIATHQKRDQYLRVQKDIIFILKDTHEBALLINRDSQVLOH 540
 DB 481 EMKRGPKPYSYLDHIIQKAIATHQKRDQYLRVQKDIIFILKDTHEBALLINRDSQVLOH 540

QY 541 KENLENNWNLIGTILKWPVNLRNYKDEQLHRVRLLYFYKPSKLYANLDDLDFAKAQ 600
 DB 541 KENLENNWNLIGTILKWPVNLRNYKDEQLHRVRLLYFYKPSKLYANLDDLDFAKAQ 600

QY 601 LTVVGQCFTEFLLESEBDGQYLEDLVKQIVOWINASSGKMPERSLQNNGLLTLSQHYF 660

DB 601 LTVVGQCFTEFLLESEBDGQYLEDLVKQIVOWINASSGKMPERSLQNNGLLTLSQHYF 660

QY 661 LFIGTILSCHPHGVKMLEKSVFQCLNLSLNQDHLKLTVSSLDYSDGLARVILSKI 720
 DB 661 LFIGTILSCHPHGVKMLEKSVFQCLNLSLNQDHLKLTVSSLDYSDGLARVILSKI 720

QY 721 LTAATDACLRYATKRLVLLRANVEFFNNWGTIELLVTQLHDKNKTITSSALDILDEACED 780
 DB 721 LTAATDACLRYATKRLVLLRANVEFFNNWGTIELLVTQLHDKNKTITSSALDILDEACED 780

QY 781 KANTHALIQMKPALSHLGDGLLLLRFLSPKGSYLNERYGVAKQLEKWHREYNKVV 840
 DB 781 KANTHALIQMKPALSHLGDGLLLLRFLSPKGSYLNERYGVAKQLEKWHREYNKVV 840

QY 841 DLIEEQLNEALTYRKPVDGDNVRRSNQRLORPHVYLPILHLYGQLVHHKTGCHLLEVN 900
 DB 841 DLIEEQLNEALTYRKPVDGDNVRRSNQRLORPHVYLPILHLYGQLVHHKTGCHLLEVN 900

QY 901 IITELCRNVRTPLDKWBEIKKLKASLWALNIGSSNWGLNLOENVIPIILKLAQCE 960
 DB 901 IITELCRNVRTPLDKWBEIKKLKASLWALNIGSSNWGLNLOENVIPIILKLAQCE 960

QY 961 VLSIRGTCVYVILGLIAKTQKQCDIILKCHNWDVRRSHKRLWPVPPDDVEQLCNELSSIPS 1020
 DB 961 VLSIRGTCVYVILGLIAKTQKQCDIILKCHNWDVRRSHKRLWPVPPDDVEQLCNELSSIPS 1020

QY 1021 TILSLNSESSTSRHNSSESVPSSMFIEDDRGSSSTSTFFLDINEDTFTFYDRSGPIK 1080
 DB 1021 TILSLNSESSTSRHNSSESVPSSMFIEDDRGSSSTSTFFLDINEDTFTFYDRSGPIK 1080

QY 1081 DKNSPFPFASSKLVKNRILNSLTLPNKHRSDDPKGKLSSESKTSNRIRITLTPEPSVD 1140
 DB 1081 DKNSPFPFASSKLVKNRILNSLTLPNKHRSDDPKGKLSSESKTSNRIRITLTPEPSVD 1140

QY 1141 FNHSDFTPISTVQTKQLTSETSMGNKHIEDTGSTPSIGENDLKFTKFNFTENHRENTSR 1200
 DB 1141 FNHSDFTPISTVQTKQLTSETSMGNKHIEDTGSTPSIGENDLKFTKFNFTENHRENTSR 1200

QY 1201 ERLVVESTSSHMKIRSQSFNTDTTSGISSMSSSPSRETGVDAITMDTDCGSMSTVVS 1260
 DB 1201 ERLVVESTSSHMKIRSQSFNTDTTSGISSMSSSPSRETGVDAITMDTDCGSMSTVVS 1260

QY 1261 TKTIKTSHVLTQSNHLSLSKSNVSLVPPGSSHTLPRAQSLKAPSIATIKSLADCNFS 1320
 DB 1261 TKTIKTSHVLTQSNHLSLSKSNVSLVPPGSSHTLPRAQSLKAPSIATIKSLADCNFS 1320

QY 1321 YTSRDAFGYATLKRLQQORMHPSLSHSEALASPAKDVLFDTDTITWKANSFESRLTPSRF 1380
 DB 1321 YTSRDAFGYATLKRLQQORMHPSLSHSEALASPAKDVLFDTDTITWKANSFESRLTPSRF 1380

QY 1381 MKALSASLKDKEILLSPINQNTLQSSSVSRVSSATYCGSDDYIGLALPVDINDIFQVK 1440
 DB 1381 MKALSASLKDKEILLSPINQNTLQSSSVSRVSSATYCGSDDYIGLALPVDINDIFQVK 1440

QY 1441 DIPYQTKNI PHDDRAGARAFADHAGGLPSGTGGLVKNSFHLRLQOQMSLITEINNSIHSDA 1500
 DB 1441 DIPYQTKNI PHDDRAGARAFADHAGGLPSGTGGLVKNSFHLRLQOQMSLITEINNSIHSDA 1500

QY 1501 SLFLESTEDTGLOEHTDDNCLVCVIEILGPOPSNOLSAICSHSDPDQDIPYSDWCQTIH 1560
 DB 1501 SLFLESTEDTGLOEHTDDNCLVCVIEILGPOPSNOLSAICSHSDPDQDIPYSDWCQTIH 1560

QY 1561 NPLEVVPKSGISGSDGVSQEGSASTSTELLIGVKTIPDDTPMCRILLRKEVLRV 1620
 DB 1561 NPLEVVPKSGISGSDGVSQEGSASTSTELLIGVKTIPDDTPMCRILLRKEVLRV 1620

QY 1621 INLSVSVTKCHETGLTITIKKYPTQFDDICLYSEVSHLLSHCTFRLPCRRIQELFQDV 1680
 DB 1621 INLSVSVTKCHETGLTITIKKYPTQFDDICLYSEVSHLLSHCTFRLPCRRIQELFQDV 1680

QY 1681 QFLQMHHEAEAVLATPPKQPIVDTSAB 1708

Db 1681 QFLQHEEAEAVLATPPKQPIVDTSAES 1708

RESULT 2

ID ADQ89852

ADQ89852 standard; protein; 1275 AA.

XX ADQ89852;

AC ADQ89852;

XX 21-OCT-2004 (first entry)

DT Antagonist of cell cycle progression polypeptide #141.

XX Cytostatic; cancer; cell division cycle; mitosis; meiosis;

DE cell cycle progression.

XX Homo sapiens.

OS WO2004063362-A2.

XX 29-JUL-2004.

XX 31-DEC-2003; 2003WO-GB005635.

XX 10-JAN-2003; 2003US-0439123P.

PR 06-MAY-2003; 2003US-0468402P.

XX (CYCL-) CYCLACEL LTD.

XX Glover D, Bell G, Frenz L, Midgley C;

XX WPI; 2004-544089/52.

DR N-PSDB; ADQ89851.

XX New cell cycle progression genes and proteins for modulating cell cycle

PT progression in cells, for preventing, treating or diagnosing cell

PT proliferative diseases (e.g. cancer) or for identifying modulators of

PT mitosis or meiosis.

XX Claim 2; SEQ ID NO 282; 461pp; English.

XX The present invention relates to a polynucleotide for preventing,

CC treating or diagnosing a disease in an individual. The composition or the

CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for

CC diagnosing, preventing or treating diseases (e.g. cell proliferative

CC diseases such as cancer) in an individual. These may also be used for

CC identifying substances capable of binding to or modulating the function

CC of the polypeptide, capable of affecting the function of the

CC corresponding gene, or capable of inhibiting the cell division cycle or

CC cell cycle progression, preferably mitosis and/or meiosis. The present

CC sequence represents an antagonist of cell cycle progression protein

CC sequence.

XX Query Match 75.3%; Score 6632; DB 8; Length 1275;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 434 ANTILPHSHSHHLCPLTLMMAASFDIPKEKRLRASAALNCLRRFHEMKRGKPKYSLH 493

Db 1 ANTILPHSHSHHLCPLTLMMAASFDIPKEKRLRASAALNCLRRFHEMKRGKPKYSLH 60

Qy 494 LDHIIQAIATHQKRDQVLRVQKDFILKQTEEALLNLRDSQVLOKHENLEWNNLIGT 553

Db 61 LDHIIQAIATHQKRDQVLRVQKDFILKQTEEALLNLRDSQVLOKHENLEWNNLIGT 120

Qy 554 ILKWPVNLRNYKDEQLHRFVRRLLYFYKPSKLYANLDDLDFAKAKQLTVVGCOFTFELL 613

Db 121 ILKWPVNLRNYKDEQLHRFVRRLLYFYKPSKLYANLDDLDFAKAKQLTVVGCOFTFELL 180

Qy 614 ESEEDGGQYLEDLVKDIVQWLNASSGMKPERSLQNNGLTTLSQHYFLFGLTSLCHPHGV 673

Db 181 ESEEDGGQYLEDLVKDIVQWLNASSGMKPERSLQNNGLTTLSQHYFLFGLTSLCHPHGV 240

Qy 674 KMLEKCSVFQCLLNLCSLKNQDHLKLTSSLDYSRDLGLARVILSKILTAATDACLRYAT 733

Db 241 KMLEKCSVFQCLLNLCSLKNQDHLKLTSSLDYSRDLGLARVILSKILTAATDACLRYAT 300

Qy 734 KHLRVLLRANVEFFNNWGIELLVTQLHDKNKTIISSEALDILDEACEDKANLHALIOMKPA 793

Db 301 KHLRVLLRANVEFFNNWGIELLVTQLHDKNKTIISSEALDILDEACEDKANLHALIOMKPA 360

Qy 794 LSHLGDGKGLLLLLFLSLIPKGFSYLNERGYAKOLEKWHREYNSKYVDLIEEQLNEALT 853

Db 361 LSHLGDGKGLLLLLFLSLIPKGFSYLNERGYAKOLEKWHREYNSKYVDLIEEQLNEALT 420

Qy 854 YRKPVGDGNYVRRSNQRLQRPVHLYPIHLYQQLVHHKGTGCHLLEVQNIITELCRNVRPD 913

Db 421 YRKPVGDGNYVRRSNQRLQRPVHLYPIHLYQQLVHHKGTGCHLLEVQNIITELCRNVRPD 480

Qy 914 LDKWEEIKKLKASLWALGNIGSSNWGLNLLQEENVIPDILKLAKOCEVLSIRGTCVYVLG 973

Db 481 LDKWEEIKKLKASLWALGNIGSSNWGLNLLQEENVIPDILKLAKOCEVLSIRGTCVYVLG 540

Qy 974 LIAKTKQCDILKCHNMDAVRHSRKLWVPVDDVEQLCNELSSIPSTLSLNESTSSRH 1033

Db 541 LIAKTKQCDILKCHNMDAVRHSRKLWVPVDDVEQLCNELSSIPSTLSLNESTSSRH 600

Qy 1034 NSESEVPSSMFILEDORFGSSSTTFPLDINEDTEPTFYDRSGPIKDKNFPFPFASKL 1093

Db 601 NSESEVPSSMFILEDORFGSSSTTFPLDINEDTEPTFYDRSGPIKDKNFPFPFASKL 660

Qy 1094 VKNRILNSLTLPNKKHRSSDPKGGKLSSEKTSNRRIRTLTEPSVDNFHSDDFTPISTV 1153

Db 661 VKNRILNSLTLPNKKHRSSDPKGGKLSSEKTSNRRIRTLTEPSVDNFHSDDFTPISTV 720

Qy 1154 QKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNFGTENHRENTSRRLVVESTSSHM 1213

Db 721 QKTLQLETSFMGNKHIEDTGSTPSIGENDLKFTKNFGTENHRENTSRRLVVESTSSHM 780

Qy 1214 KIRSQSFNTDTTSGISSMSSPSRETGVDAITMDTCGSMSTVSVTKTITKSHYLPQ 1273

Db 781 KIRSQSFNTDTTSGISSMSSPSRETGVDAITMDTCGSMSTVSVTKTITKSHYLPQ 840

Qy 1274 SNHLSLKSNSVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFSYTSRDAFGYATL 1333

Db 841 SNHLSLKSNSVSLVPPGSSHTLPRRAQSLKAPSIATIKSLADCNFSYTSRDAFGYATL 900

Qy 1334 KRLQOQRMHPSLSHSEALASPAKDVLTDTITMKANSFESRLTPSRFMKALSASLDKED 1393

Db 901 KRLQOQRMHPSLSHSEALASPAKDVLTDTITMKANSFESRLTPSRFMKALSASLDKED 960

Qy 1394 LLSPINQNTLQRSSSVRSVSSSATYGGSDDYIGLALPVDINDIFOVKDIPYFQTKNIIPP 1453

Db 961 LLSPINQNTLQRSSSVRSVSSSATYGGSDDYIGLALPVDINDIFOVKDIPYFQTKNIIPP 1020

Qy 1454 DDRCARAFADHAGGLPGSTGGLVKNFHLRQOQSLTEIMNSIHSDASLFLESTEDTGLQ 1513

Db 1021 DDRCARAFADHAGGLPGSTGGLVKNFHLRQOQSLTEIMNSIHSDASLFLESTEDTGLQ 1080

Qy 1514 EHTDDNCLYCVIBILFQPSNQLSAICSHSDPDODIPYSDWCBOQTINHPLEVPVSKFSGI 1573

Db 1081 EHTDDNCLYCVIBILFQPSNQLSAICSHSDPDODIPYSDWCBOQTINHPLEVPVSKFSGI 1140

Qy 1574 SGCSGDSGVQEGSASSTKSTELLGKVTIPDTPMCRILLRKEVLRLVNLSSSVSTKCHE 1633

Db 1141 SGCSGDSGVQEGSASSTKSTELLGKVTIPDTPMCRILLRKEVLRLVNLSSSVSTKCHE 1200

Qy 1634 TGLLTIKKYPQTPDDICLYSEVSHLSHCTFRFLPCRRFIQELFQDVQFLQHEEAEAVL 1693

Db 1201 TGLLTIKKYPQTPDDICLYSEVSHLSHCTFRFLPCRRFIQELFQDVQFLQHEEAEAVL 1260

Qy 1694 ATPPKQPIVDTSAES 1708

Db 1261 ATPPKQPIVDTSAES 1275

RESULT 3

AD71253
ID ADE71253 standard; protein; 791 AA.

XX AC ADE71253;

XX 29-JAN-2004 (first entry)

DT Novel human protein #7.

DE human; novel protein; drug.

XX KW Homo sapiens.

XX OS JP2002345493-A.

XX PN 03-DEC-2002.

XX 29-MAR-2001; 2002JP-00049046.

XX PF 29-MAR-2001; 2001JP-00095524.

XX PR (KAZU-) ZH KAZUSA DNA KENKYUSHO.

XX PA WPI; 2003-460885/44.

XX DR N-PSDB; ADE71191.

XX PT A gene and a protein encoded by it, used in drugs.

XX PS Disclosure; Page 52-55; 257pp; Japanese.

XX CC The invention comprises the amino acid and coding sequences of novel human proteins. The DNA and protein sequences of the invention are used in drugs. The present amino acid sequence represents a novel human protein of the invention.

XX SQ Sequence 791 AA;

Query Match 45.6%; Score 4014; DB 7; Length 791;

Best Local Similarity 100.0%; Pred. No. 2.9e-314;

Matches 778; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	931	GNIGSSNWGLNLQENVIPIIDILKAKQCEVLSIRGTCVVVLGLIAKTKQGGDILKCHRW	990
DB	14	GNIGSSNWGLNLQENVIPIIDILKAKQCEVLSIRGTCVVVLGLIAKTKQGGDILKCHRW	73
QY	991	DAVHRSKHLWPVVPDDVEQLCNELSSIPSTLSNSESTSSRHNSSESVPSMFLEDD	1050
DB	74	DAVHRSKHLWPVVPDDVEQLCNELSSIPSTLSNSESTSSRHNSSESVPSMFLEDD	133
QY	1051	RGSSSTSTFFLDINEDTEPTFYDRSGPIKDNKSPFPFASCKLVKRIINSLTLPNKKHR	1110
DB	134	RGSSSTSTFFLDINEDTEPTFYDRSGPIKDNKSPFPFASCKLVKRIINSLTLPNKKHR	193
QY	1111	SSSDPKGKLSSESKTSNRRIITLTPSDVDNHSDDFTPISTVQKTLQLETSPMGKHIE	1170
DB	194	SSSDPKGKLSSESKTSNRRIITLTPSDVDNHSDDFTPISTVQKTLQLETSPMGKHIE	253
QY	1171	DTGSTPSIGENDLKFTKNFGTENHRENTSRERLIVVESSTSSHMKIRSQSFNTDTTSGIS	1230
DB	254	DTGSTPSIGENDLKFTKNFGTENHRENTSRERLIVVESSTSSHMKIRSQSFNTDTTSGIS	313
QY	1231	SNSSSPSRRTVGVDATMTDCGSMSTVSTTKITKSHYLTQSNHLSKNSVSLVPP	1290
DB	314	SNSSSPSRRTVGVDATMTDCGSMSTVSTTKITKSHYLTQSNHLSKNSVSLVPP	373
QY	1291	GSSHTLPRAQSLKAPSIATIKSLADCNFSYSSRDAGFATLKRLOQORMHPSLSHSEA	1350
DB	374	GSSHTLPRAQSLKAPSIATIKSLADCNFSYSSRDAGFATLKRLOQORMHPSLSHSEA	433
QY	1351	LASPAKDVLTFTITWKANSFESRLTPSRFMKALSYASLDKEDLLSPINQNTLQRSSVR	1410

DB	434	LASPAKDVLTFTITWKANSFESRLTPSRFMKALSYASLDKEDLLSPINQNTLQRSSVR	493
QY	1411	SMVSSATYGGSDYIGIALPVDINDIFOVKDIPYFQTKNIPPHDDRGARAFADAGLPS	1470
DB	494	SMVSSATYGGSDYIGIALPVDINDIFOVKDIPYFQTKNIPPHDDRGARAFADAGLPS	553
QY	1471	GTGGLVNSFHLLRQOMSLTEIMNSIHSASLFLBSTEDEGLQEHDDNCLYCVCIIEILG	1530
DB	554	GTGGLVNSFHLLRQOMSLTEIMNSIHSASLFLBSTEDEGLQEHDDNCLYCVCIIEILG	613
QY	1531	FQPSNQLSAICSHSDFQDIPYSDWCETIHNPLEVVPSPKSGISGSDGVSQSGSASTK	1590
DB	614	FQPSNQLSAICSHSDFQDIPYSDWCETIHNPLEVVPSPKSGISGSDGVSQSGSASTK	673
QY	1591	STELLGVKTIIPDDTPMCRILLRKEVLRVLVNLSSSVTKCHETGLTTKEKYPQTFDDI	1650
DB	674	STELLGVKTIIPDDTPMCRILLRKEVLRVLVNLSSSVTKCHETGLTTKEKYPQTFDDI	733
QY	1651	CLYSEVSHLLSHCTPRLPCRRIQELFQDVQFLQMHHEAEAVLATPPKQPIVDTSAES	1708
DB	734	CLYSEVSHLLSHCTPRLPCRRIQELFQDVQFLQMHHEAEAVLATPPKQPIVDTSAES	791

RESULT 4

ADC31052

ID ADC31052 standard; protein; 336 AA.

XX AC ADC31052;

XX 18-DEC-2003 (first entry)

XX Human novel polypeptide sequence, SEQ ID NO:1134.

XX Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnerary;
 KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 5.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

XX Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

XX N-PSDB; ADC30081.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

XX treating conditions such as neurodegenerative diseases, anemias, platelet

XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

XX cancer.

XX Claim 20; SEQ ID NO 1134; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-

XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

XX invention also relates to nucleic acid sequences over 99% identical with

XX the novel human cDNAs. The invention additionally encompasses expression

Db 366 SNGFVAEGRSILPTLAARAPSVVQHILALMLYCFLETGLLNALVVEVVSSDQFVSQVAT 425
Qy 425 ILLGELLHANTILPHSHSHHILCLPTLMNMAASFDIPKEKRLRASAALNCLRPHEMKK 484
Db 426 VLIIGKILQMLHPTLPPDICTSPALPTLVSHATL-----GNOQANAFAAVALQNYOKLLR 479
Qy 485 RQPKPYSLLHLDHIQ--KAIATHQRKDQVLRVKQD--IFILKDTFEE-----526
Db 480 QRPASRSLFDSIIQGGALIQTLRFR--RLHNVQEQAGPVLQLOETAEAAAFVPPVPTQFR 538
Qy 527 -----ALINLR-----DSQVLQHK 541
Db 539 GTLDRRLDSVSSDESNSQASTSSRSFRLKXKLPQALYDNPFAFNRLLTDSRLVLSQA 598
Qy 542 ENLEWNWNLIGTILKWPVNLRNYKDEQLHRFVRRLLYPKSSKLYANLDDPAKAKQL 601
Db 599 DAHLMDWDVITILK--SNLRKLDYTOGKELKRLVDYKPRNRFHQDL--VPGQRL 653
Qy 602 TV---VGCQFTFELLESE--DGOGYLEDLVKDIVQWLNA--SSGMPKPSRLQNNGLLTTL 655
Db 654 PTYVSAGLDLIDLSSNELECMRFITDYFSDISQOLAAVTTNSRAHDCLPSPQHNWNTM 713
Qy 656 SQHYFLFIGTSLCHPHGVKMLKCSVFQCLLMCLSKNQDHLKLTVSSLDYSRDLARV 715
Db 714 CQOYFLYIGMCHTVKGIKLVKNTTVFVYLINLRVTDHVCYVKKLIVSGLNYSYEKLPRQ 773
Qy 716 ILSKILTAADACRLVATKHLRVLRLANVEFFNNGWIELLVTLQHLKNTKTSISEALDIID 775
Db 774 VLEKALTSAKTRGLYSTOPMAYLLRPLRPFVFWGPIIINQTRDSRVSGLAAMDVLE 833
Qy 776 EACEDKANLHALQMPKALSHLDGKGLILLRLFLSPKGFSLNNGRYVAKOLEKWHREY 835
Db 834 EACDKYILEEIVSRPNLTHRGDAGRLLMARYSLPRGLNSTMAR--TEDEIRYRWNY 891
Qy 836 NSKYVDLIEQLNEALTTRKPYVDGNYVRNSRQRLQRPVLYPIHLYQOLVHHKTCGHL 895
Db 892 NKRYVLLVEADTHSSLTHIRNEDGYVSRNCNRPQTVPPNVAPHLYGMAQTGQGMTA 951
Qy 896 LEVONIITELCRNVTRPDLKWEIKKILKASLWALNGISSNWGLNLOEEN--VIPDIL 953
Db 952 LRKHGDLPOLLELRRAKCTDDAECELEKAAIWAHAHASTHNGIEYFVELNARLYEKLI 1011
Qy 954 KLAQCEVLSIRGCVVVLGLIAKTQGCIDILKHNWDVAVRHGRKHLWPV--VPDDVEQLC 1012
Db 1012 VLVTKCEVSVRATCPFSAAGLIAGTQAGANILFKLNWLSVRHDKNTMWPVHQPED-----1066
Qy 1013 NELSSIPSTLSLNESTSSRHNSSESVPSSMFILEDD---RFGSSSTSTFFLDINEDIE 1069
Db 1067 -----WMSSQYTPVRH--YYEDVPADNMTMDDYIERFYETG--SDFWMLADQT- 1112
Qy 1070 PTFYDRSGPIKDKNSFPFPAASSKLAVKNRILNLSLTPNKKHRSSSDPKGGKLSSESKTSNR 1129
Db 1113 -----DASG-----GGGVAGTVVG-----MEPIVQDFNATITITDS-----1144
Qy 1130 RIRLTPEPSVDFNHSDDFTP-----ISTVQKTLOLETSPMGKXH---IEDTGSTPSIGEN 1181
Db 1145 -VRIT-----DDVHPRMLVGVAASKTLPGLSNLQKQKHORSLSKTTDVIS-- 1192
Qy 1182 DLKFTKXPGTENHRENTSRERLVSESSTSHMKIRSQSFNTDTTISGISMSSSPSRETV 1241
Db 1193 --LLGSGVGTMSG-----AVLYPTPYQHRIRYNSC--TDSNTSGVSSCESVTGTAA 1240
Qy 1242 GVDATMTDTCGSMSTVVTXTIKTSHYLTQSNHLSLSKNSVSLVPPGSSHTLPRAQ 1301
Db 1241 AYAAANLEQFPPLSPISPMENL-----EDELFRNQLATSMLP-----1281
Qy 1302 SLKAPSIATKSLADCNFSYTSRDAFGYATLKLOQRMHPISLSHEALASPAKVLF 1361
Db 1282 -----STLSPMA-----MKGYVQLRSL--RKHSRPVFSSESAEYDFAEILDT 1322
Qy 1362 DTITWKANSFRSLTPSRFMKALSASLDKE--DL---LSPINQNTLQRSSSVRSVSSA 1416
Db 1323 PEVQMR-----KLDWTHSHRRLKVRSLDRQLSDVYRRLSADENLVLLTINAPKFFPPN 1376

Qy 1417 TYGSSDDYIGLALPVDINDIFQVKDIP-YFQTKNIPPHDDRGARAFADAGGLPSGTGGL 1475
Db 1377 DLKG-PCYAGICLPKXVLDLFPFNLNLSRTVVSRIQDQDIVGI-----1418
Qy 1476 VKNSFHLRQOMSGITEIMNSIHSDASLFLESTEDTG-----LOE---HTDDNCLYC 1523
Db 1419 --NLNTMLRPOCLNDSLNTNEGDESSVSSLSVSSASRQTRLOQQQAKHSRSLCLHC 1476
Qy 1524 VCIEILFQPSNQLSALCSHSD---FQDIP---YSDWCEQTIHNPLEVVP--SKFSGIS 1574
Db 1477 -----ARSPRQQRNDSGSHSGGGGASLAPCELYSSAAALVAAGIQAGFVLAKKSGSS 1530
Qy 1575 GCSGVQOE--GSASSTKSTELLGVKTPDDTPMCRILLRKEVLRVLVINLSSSVSTKCHE 1633
Db 1531 --AQGSAALGADISFHSPEMSLSESLPD-----RLTASILYNVORLANPVSAKQSK 1581
Qy 1634 TGLTITKEKVPQTPDDTCLYSEVSHLLSHCTFRPLPCRFRFOELFQDVQF-----1682
Db 1582 MALLEKQKHPAFQDICYSEACKTIGRSSYRMIARFLQELFLDLNFDLSFYVEPOLII 1641
Qy 1683 ---LQMHHEEAVALAT-----PPKQPIV 1702
Db 1642 GARKFSAEKAABATAPTAMQAMPFKQIIL 1670

RESULT 6
ABB63631
ID ABB63631 standard; protein; 1922 AA.
XX
AC ABB63631;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 17685.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
PI WPI; 2001-656860/75.
XX
DR N-PSDB; ABL07734.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
XX
PS Disclosure; SEQ ID NO 17685; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX	SQ	Sequence 1922 AA;	Query Match	18.8%;	Score 1660;	DB 4;	Length 1922;
			Best Local Similarity	28.0%;	Pred. No. 1.8e-123;		
			Matches 509;	Conservative 314;	Mismatches 684;	Indels 314;	Gaps 54;
QY	27	PLDLTRBSSDNLREILQWVAR-----LQGVSNMRKLGHLNNFTKLLCDIR--GHS	73				
DB	5	PLDLSSQDEDFVRLDPQSAANAPEIYSMLCBEETRTDKRLFLINALASLCLGARKGSH	64				
QY	74	EEKGLGFHYEDIIICLRALLANEKVEVRAGLRALRYLIQDSSILQKVLKLVKDVYLIARCI	133				
DB	65	HSNRFTTEELLYCLASLSVHTFTQVRAAALRTIRYAISTFNDIKTFNALQLOHLKCRSI	124				
QY	134	DIQOSNEVERTQALRVLKRMITVNASLPSSVWTNSLIAVGDNGDQERDMVRACIAIICE	193				
DB	125	DLMLKNDDERVOALKLVKMLAIAPEDISPVVRCVLSLAUGSIEENDNLIRACIATLAE	184				
QY	194	LALQNPVEVALRGGLNTILKNVDCQLSRINEALITITILHLNHPKTRQYVRADVELERI	253				
DB	185	FAVLNPAIIIVCGGVTSITRNVLCEHNPRIAESLCGVLLYLEWPTQRI--CGVRLDCL	242				
QY	254	LAPYTDHYRHSPTAEQKXEDREARFLASKMGIIATFRSMAGIINLCKPGN--SGTQSL	312				
DB	243	AAPYCDFTYQS---IVDNKDARELRYTSCRLALLSVLRSWTGLTBECDSPKSGLKAI	299				
QY	313	IGVLCIPNMEIRRGILLEVLYDI FRPLPVPVTEETFEALLSVDPGRFODSWMRLSDGFVAAE	372				
DB	300	VDALYLNQIEVRKAILDLLEYELLTLTPQTWTDVVAVALQAVDPDSFDQTWLLSNGFVSAE	359				
QY	373	AKTILPHRARSRPDMNYLALIIISAFIRNGLLEGLVEVITNSDDHISVRATILGLBLH	432				
DB	360	GRSILPTLAARAPSVVEQHLALMLYCFLETGLLNALVEVVVSSDOFVSQVATVLIGKILQ	419				
QY	433	MANTILPHSHSHLLHCLPTLMMMAASPDIPKEKLRASAAALNCLKRHEMKRGKPKPSYL	492				
DB	420	LMHTLPPDICCTSPALPTLVSHATL-----GNQQAANAVALQNTQKLLRQRPASRSL	473				
QY	493	HLDHIIQ--KAIAITHQKRDQVLRVQKD--IFILKDTTEE-----	526				
DB	474	FLDSIIQGGALIQTLFR--RHLNVQEQAGPVLIQETAEAAAFVAVPVTQPRGTLDRSRL	532				
QY	527	-----ALLINLR-----DSQVLQHKENLEWNWN	549				
DB	533	DSVSSSDSNSQASTSSRSRSPRLKPKFLPQALYDNFRAFNRLITDSRVLSQADAHLDWD	592				
QY	550	LIGTILKWPVNLRYKDBQLHFRVRLLYPKSSKLYANLDDFAKAKQLTV---VGC	606				
DB	593	VITILK---SNLRKLDYTGKFLKRLVDFYKPRNNRFSHQDL--VPGQRLPITYSAGL	647				
QY	607	QTFBFLLESEB--DCGGVLEDLVKDIIQVWLNA--SSGMKPRSLQNNGLLITLSQHYFLFI	663				
DB	648	DLIDVLSNSELECMRFITDYFSDI SQQLAAVTTNSRAHDCLPFQHMNTMCOQYFLYI	707				
QY	664	GTLSCPHGVKMLKBCSVFQCLLNLCSLKNQDHLLKLTVSLSYSDRDLARVILSKILTA	723				
DB	708	GRMCKTVKGI EVLKNVTTFEYLINLRVTDHVCVYVKLIVSGLNYSYEKLPRQVLEKALTS	767				
QY	724	ATDACRLYATKHLRVILRANVEFPNNWGIELLVTLQHDKNTKTISSDALDIDDEACEDKAN	783				
DB	768	AKTRGLYSTQFMVALRLARLPHEFVWGIPIIINQTRDSRDSVSGLAAMDVLEACHDKYY	827				
QY	784	LHALIQMKPALSHLDGKGLLALLRFLSPIKGFSVLNERGVYVAKOLEKWHREYNSKYVDLI	843				
DB	828	LEEIVSRWPNLTHRGDAGRLMARIYSLSPRGNTMAR--IEDIRWRNGYKRYVLLV	895				
QY	844	BEQLENEALTTYRKPEVDGNTVYRRSNQRQLRPHVYLPILHYLQGLVHHHTGCHLLEVQNIIT	903				
DB	886	EADTHSSILTHIRNEDGYISRENCNQRQPTVPPNVAAPHLYGMAQTQCGMTALRKHGDL	945				
QY	904	ELCRNVRTPLDKWEEIKKULASLWALGNTGSSNGLNLQEEV--VIPDILKLAKQCEV	961				

D	b	946	QLLELLRRAKTDDAECLCAAIAWALAHASTHSNGIEYFVELNARLYEKLIVLVTKEV	1000
Q	y	962	LSIRGTCVYVLGIATKQGCDILKCHNWDAVRHRSKHLWPV-VPDDEVQLCNELSSIPS	1020
D	b	1006	YSVRATCFESALGLIAGTOAGANILFKLNLSVLRHDKNMPPVHQEPD-	1052
Q	y	1021	TLSLNSESTSRHRNSESESVPSSFMLEDD--RFGSSTSTFFLDINEDTEPTFYDRSG	1077
D	b	1053	--WMSSQVTPVRH--YYEDVPADMTMMDDVIERYFETG-SDFWHMLADQT----DASG	1102
Q	y	1078	PIKDKNFPFPASSKLVKRIILNSLTPLNKHKRSSDPKGGKLSSESKTSNRRIRLTEP	1137
D	b	1103	-----GGGVAGTVVG-----MHPIVQDPNATIITTTDS-----VRTTT--	1135
Q	y	1138	SVDFNHSDFTP-----ISTVQTKLOLETSMGNKH---IEDTGSPISIGENDLKFTKNF	1189
D	b	1136	-----DDVHPRMLVCVAASKTLPEGSNLROCKQRSISESKTTDVIS-----LLGSGV	1184
Q	y	1190	GTEMHRENTGRERLWVESSTSSHMKIRSOFNTDTTTSIGISSMSSPSRETGVGDATTMD	1249
D	b	1185	GTMSG-----AVLYPTPYQHRIYRNSC-TDSNTSGVSSCESVTGRTAAAYAAAAANE	1234
Q	y	1250	TDCGSMVTYSTTKITSHVLTQSNHLSLSKNSVSLVPPGSSHLLPRAAQSCLKAPSTA	1309
D	b	1235	LQQPLSPIFSMNLL-----ESDELFRNQLATSMLP-----S	1268
Q	y	1310	TIKSLADCNFYSTSRDAFYATLKKLQOORMHPSLSHSEALASPDKDLVTDITTMKAN	1369
D	b	1269	TLSPMA-----MGKYVOLRSJ-RKHSRPVFESASEPYDPAEILDTPEQMR--	1314
Q	y	1370	SFESRLTPSRFMKALSAYSADKE--DL---LSPINQNTLQRSSSVRSMVSATYGGSDDY	1424
D	b	1315	---KLDWTHSHRRLKYRSLDRQLSDYVRLSADELNVLPLTNAPKFLPPNDLKG-PCY	1369
Q	y	1425	IGIALPVDINDIIFOVKDIP-YFOTKNI PHDDRAGARAFAHDAGGLPSGTGGLVKSFHLL	1483
D	b	1370	AGICLPKNVLDPPTTRLNLSRYVSRDIQODIVGI-----NILNTM	1410
Q	y	1484	RQOWSLTEIMNSTHSDASFLESTEDTG-----LQE---HTDNCLCYVCIBILGF	1531
D	b	1411	LRQCUNDLSITNCGGHESSVISLSVSSASRRQTRLOOGAKHSRLCLHC-----AR	1464
Q	y	1532	QPNQLSAICSHSD----FDIP---YDWCEQTIHNPLEVVP--SKFSGISGCSGVSQ	1582
D	b	1465	SPQORNDSGSHNGGGGASLAPCELYSSAAAALVAAGIQAGVPLAKKSGSS--AQGASA	1522
Q	y	1583	E-GSASTKSTELLGVKTIPDPTPMCRILLRKVEURLVINLSSSVSTKCHETGLLTIKE	1641
D	b	1523	ALGADISFHSPEMLSSEDLPD-----RLTASILYNQVRLANPVSAKQSMALLLEKQ	1575
Q	y	1642	KYPOTPDICLYSEVSHLLSCHCTFRLPCRFTQELPODVQF-----LQMHW	1686
D	b	1576	KHPHAQDICLYSEACKTGRSSYMIARRFIQELFLDLNFDSFVYVEPOLIIGARKFSAB	1635
Q	y	1687	EEBAEAVLAT-----PPKQPPIV	1702
D	b	1636	EKAEAAPTAMQAMPFKQIIL	1656
RESULT 7				
A	D	C49296	ADC49296	
I	D	C49296	standard; protein; 329 AA.	
X	X	C49296		
A	C	C49296		
X	X	C49296		
D	T	C49296	18-DEC-2003 (first entry)	
X	X	C49296		
D	E	C49296	Polypeptide-cell regulation factor -36.19 protein.	
X	X	C49296		
K	W	C49296	Polypeptide-cell regulation factor -36.19; enzyme; cancer; tumour;	
K	W	C49296	haemopathy; human immunodeficiency virus; HIV; immunological disease;	
X	X	C49296	inflammation; antagonist.	

Polypeptide-cell regulation factor -36.19 protein.

KW Polypeptide-cell regulation factor -36.19; enzyme; cancer; tumour;
KW haemopathy; human immunodeficiency virus; HIV; immunological disease;
KW inflammation; antagonist.

OS Unidentified.
XX
PN CN1382698-A.
XX
PD 04-DEC-2002.
XX
XX 26-APR-2001; 2001CN-00112725.
XX
PR 26-APR-2001; 2001CN-00112725.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2003-269462/27.
XX
XX N-PSDB; ADC49295.
PT Polypeptide-cell regulation factor-36.19 and polynucleotide for coding
PT it.
XX
XX Claim 1; Page 26-27; 31pp; Chinese.
XX
XX The present invention discloses a novel polypeptide-cell regulation
CC factor -36.19, polynucleotide coding for the polypeptide and method for
CC producing this polypeptide by using DNA recombination technology. The
CC invention also discloses the method for curing several diseases, such as
CC cancer, malignant tumour, haemopathy, human immunodeficiency virus (HIV)
CC infection, immunological diseases and various inflammations by using the
CC said polypeptide. The invention also discloses an antagonist for resisting
CC said polypeptide and its therapeutic action and also discloses the
CC application of the polynucleotide for coding this novel Polypeptide-cell
CC regulation factor -36.19. The sequence presented is the polypeptide-cell
CC regulation factor -36.19 protein.
XX
SQ Sequence 329 AA;
Query Match 18.5%; Score 1631; DB 7; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.9e-122;
Matches 324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAAIGRSLKRLVRGRNDSGEENVPLDLTREPNDLREILQNVARLQGVSNMRKLGHL 60
DB 1 MAAIGRSLKRLVRGRNDSGEENVPLDLTREPNDLREILQNVARLQGVSNMRKLGHL 60
QY 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRALLNEAKEVRAAGLRALRYLIQDSSILQKV 120
DB 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRALLNEAKEVRAAGLRALRYLIQDSSILQKV 120
QY 121 LKLVYDLIARCIDIQOSNEVERTQALRVKMTITVNASLFPSSVTNSLIAGVNDGLQER 180
DB 121 LKLVYDLIARCIDIQOSNEVERTQALRVKMTITVNASLFPSSVTNSLIAGVNDGLQER 180
QY 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITITLHLNHPKT 240
DB 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITITLHLNHPKT 240
QY 241 RQVRAVLELERILAPYTFDHRHSPTDAGQLKEDREARFLASKMGIIATFRSWAGIIN 300
DB 241 RQVRAVLELERILAPYTFDHRHSPTDAGQLKEDREARFLASKMGIIATFRSWAGIIN 300
QY 301 LCKPFGNSGIGSLIGVLCIPNWEIRRG 326
DB 301 LCKPFGNSGIGSLIGVLCIPNWEIRVG 326
RESULT 8
ADRI0059
ID ADRI0059 standard; protein; 313 AA.
XX
AC ADRI0059;
XX
XX 04-NOV-2004 (first entry)
XX

DE Human protein useful for treating neurological disease Seq 3565.
XX
KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
XX EP1447413-A2.
PN
XX 18-AUG-2004.
PD
XX 12-FEB-2004; 2004EP-00003145.
PF
XX 14-FEB-2003; 2003JP-00102207.
PR
XX 09-MAY-2003; 2003JP-00131452.
PR
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI; 2004-583265/57.
DR N-PSDB; ADR08103.
DR
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 3565; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC office.
XX
SQ Sequence 313 AA;
Query Match 17.4%; Score 1535; DB 8; Length 313;
Best Local Similarity 94.8%; Pred. No. 1e-114;
Matches 309; Conservative 0; Mismatches 1; Indels 16; Gaps 1;
QY 1 MAAIGRSLKRLVRGRNDSGEENVPLDLTREPNDLREILQNVARLQGVSNMRKLGHL 60
DB 1 MAAIGRSLKRLV-----REPSDNLREILQNVARLQGVSNMRKLGHL 44
QY 61 NNFTKLLCDIGHSEKLGPHYEDIIICLRALLNEAKEVRAAGLRALRYLIQDSSILQKV 120
DB 45 NNFTKLLCDIGHSEKLGPHYEDIIICLRALLNEAKEVRAAGLRALRYLIQDSSILQKV 104
QY 121 LKLVYDLIARCIDIQOSNEVERTQALRVKMTITVNASLFPSSVTNSLIAGVNDGLQER 180
DB 105 LKLVYDLIARCIDIQOSNEVERTQALRVKMTITVNASLFPSSVTNSLIAGVNDGLQER 164
QY 181 DRMVRACIAIICELALQNPVVALRGGLNTILKNVIDCOLSRINEALITITLHLNHPKT 240
|||

Db 165 DMVRACIAIICELALQNPVEVALRGGLNTILKNVICQLSRINEALITLILHLNHPKT 224
Qy 241 RQYVRADVELERILAPYTDHYRHSPTAEQQLKEDREARFLASKMGIIATFRSWAGIIN 300
Db 225 RQYVRADVELERILAPYTDHYRHSPTAEQQLKEDREARFLASKMGIIATFRSWAGIIN 284
Qy 301 LCKPQNGSIQSLIGVLICIPNMEIRRG 326
Db 285 LCKPQNGSIQSLIGVLICIPNMEIRVG 310

RESULT 9
ABG01058
ID ABG01058 standard; protein; 394 AA.
XX AC ABG01058;
XX XX
DT 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #1049.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS65245.
XX XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 31417; 103pp; English.
XX PS
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 394 AA;

Query Match 12.8%; Score 1127; DB 4; Length 394;
Best Local Similarity 84.9%; Pred. No. 1.4e-81;
Matches 220; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
Qy 1380 FMKALSVASLDKEDLLSPINQNTLQSSSVRSVMVSSATYGGSDYIGLALPVDINDIFQV 1439
Db 102 FMKALSVASLDKEDLLSPINQNTLQSSSVRSVMVSSATYGGSDYIGLALPVDINDIFQV 161
Qy 1440 KDIPYFQTKNIPPHDDRGARAFADAGGLPGTGLVKNFHLIRQOMSLTEIMNSIHS 1499
Db 162 KDIPYFQTKNIPPHDDRGARAFADAG----- 188
Qy 1500 ASLFLESTEDTGLQEHDDNCLYCVCIETILGPOPSNOLSAICSHSDPDQIPYSDWCQTI 1559
Db 189 -----ESTEDTGLQEHDDNCLYCVCIDILGFOPSNOLSAICSHSDPDQIPYSDWCQTI 243
Qy 1560 HNPLEVVPKFSGICSGDGVSGSASSTKSTELLIGVKTIPDDTPMCRILLRKEVRL 1619
Db 244 HNPLEVVPKFSGICSGDGVSGSASSTKSTELLIGVKTIPDDTPMCRILLRKEVRL 303
Qy 1620 VINLSSSVSTKCHETGLLT 1638
Db 304 VINLSSSVSTKCHETGLLT 322

RESULT 10
ABG16310
ID ABG16310 standard; protein; 394 AA.
XX AC ABG16310;
XX XX
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #16301.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS80497.
XX XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 46669; 103pp; English.
XX PS
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging

of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 394 AA;

Query Match 12.8%; Score 1127; DB 4; Length 394;
Best Local Similarity 84.9%; Pred. No. 1.4e-81;
Matches 220; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
QY 1380 FMKALSYASLDKEDLLSPINQNTLQSSSVSRVSSVSSATYGGSDDYIGLALPVDINDIFQV 1439
Db 102 FMKALSYASLDKEDLLSPINQNTLQSSSVSRVSSVSSATYGGSDDYIGLALPVDINDIFQV 161
QY 1440 KDIPFQTKNIPPHDDRGARAFADHAGLPGSTGGLVKNFSHLLRQOMSLTMSIHSD 1499
Db 162 KDIPFQTKNIPPHDDRGARAFADHAG----- 188
QY 1500 ASLFLESTEDTGLQHTDDNCLYCVICILGFPQSNQLSAICSHSDPDIYSWCEQTI 1559
Db 189 -----ESTEDTGLQHTDDNCLYCVICILGFPQSNQLSAICSHSDPDIYSWCEQTI 243
QY 1560 HNPLEVPSKFSIGSCSDGVSGSASSTKSTELLGVKTIPTDPTMCRILLRKEVLRL 1619
Db 244 HNPLEVPSKFSIGSCSDGVSGSASSTKSTELLGVKTIPTDPTMCRILLRKEVLRL 303
QY 1620 VINLSSSVSTKCHETGLLT 1638
Db 304 VINLSSSVSTKCHETGLLT 322

RESULT 11

AD32769
ID ADC32769 standard; protein; 394 AA.

XX AC ADC32769;

XX DT 18-DEC-2003 (first entry)

XX DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:2851.

XX KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 5.

XX OS Homo sapiens.

XX PN WO2003029271-A2.

XX PD 10-APR-2003.

XX PF 24-SEP-2002; 2002WO-US030474.

XX PR 24-SEP-2001; 2001US-0324631P.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang TX, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;

XX DR WPI; 2003-371981/35.

XX N-PSDB; ADC32002.

XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.

XX PS Example 2; SEQ ID NO 2851; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig-
CC encoded polypeptide sequence used in an example of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 394 AA;

Query Match 12.8%; Score 1127; DB 7; Length 394;

Best Local Similarity 84.9%; Pred. No. 1.4e-81;

Matches 220; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1380 FMKALSYASLDKEDLLSPINQNTLQSSSVSRVSSVSSATYGGSDDYIGLALPVDINDIFQV 1439
Db 102 FMKALSYASLDKEDLLSPINQNTLQSSSVSRVSSVSSATYGGSDDYIGLALPVDINDIFQV 161
QY 1440 KDIPFQTKNIPPHDDRGARAFADHAGLPGSTGGLVKNFSHLLRQOMSLTMSIHSD 1499
Db 162 KDIPFQTKNIPPHDDRGARAFADHAG----- 188

QY 1500 ASLFLESTEDTGLQHTDDNCLYCVICILGFPQSNQLSAICSHSDPDIYSWCEQTI 1559
Db 189 -----ESTEDTGLQHTDDNCLYCVICILGFPQSNQLSAICSHSDPDIYSWCEQTI 243

QY 1560 HNPLEVPSKFSIGSCSDGVSGSASSTKSTELLGVKTIPTDPTMCRILLRKEVLRL 1619
Db 244 HNPLEVPSKFSIGSCSDGVSGSASSTKSTELLGVKTIPTDPTMCRILLRKEVLRL 303

QY 1620 VINLSSSVSTKCHETGLLT 1638

Db 304 VINLSSSVSTKCHETGLLT 322

RESULT 12
ABR53386

ADK63324;
06-MAY-2004 (first entry)
Disease treating protein complex-derived protein #991.
protein complex; drug target; diagnosis.
Unidentified.
EP1338608-A2.
27-AUG-2003.
20-DEC-2002; 2002EP-00102902.
20-DEC-2001; 2001EP-00130253.
(CELL-) CELLZOME AG.
Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;
Marzioch M, Grandi F, Krause R, Kruse U, Merino A, Bauch A;
Michon A, Leutwein C, Rick J;
WPI; 2003-638460/61.
N-PSDB; ADK63325.
New proteins and protein complexes from eukaryotes, useful as targets in
drug screening, or in diagnosing or screening for the presence of a
disease or disorder, or a predisposition for developing a disease or
disorder in a subject.
Disclosure; SEQ ID NO 1981; 13pp; English.
The invention relates to novel protein complexes comprising a first and a
second protein, or its derivative, fragment, homologue or variant. The
proteins are selected from given protein complexes, which are not defined
in the specification. The variants are encoded by nucleic acids that
hybridize to the nucleic acids encoding the proteins under low stringency
conditions. The protein complexes are useful as targets for an active
agent of a pharmaceutical. These protein complexes are particularly
useful as drug targets for the treatment or preventing of a disease or
disorder. The complexes and methods above are useful in diagnosing or
screening for the presence of a disease or disorder or a predisposition
for developing a disease or disorder in a subject. These are also useful
in screening for a drug for treatment or prevention of a disease or
disorder. The molecule that modulates the amount, activity or protein
components of the complex is useful for the manufacture of a medicament
for the treatment or prevention of a disease or disorder. This sequence
corresponds to a protein of the invention. (Note: the sequence data for
this patent did not form part of the printed specification but was
obtained from the EPO in electronic format).
Sequence 1430 AA;
Query Match 8.2%; Score 720.5; DB 7; Length 1430;
Best Local Similarity 23.1%; Pred No. 8.3e-48;
Matches 301; Conservative 218; Mismatches 531; Indels 255; Gaps 39;
QY 18 RNDGSEEN-----VPLD-----LTREPSDNLREILQNVAR----- 47
DB 203 RNDQGNDSBELGDHDSLTQVTLDDVLTPTSGTERNQQNLNRNSTVNRNENHS 262
QY 48 -----LQGVNM-----RKLGLHNN-----FTKLLCDIG----- 71
DB 263 TLSIPDLGSKNVNLTGTEDKLDGLENENQIFTSITTEAATLWLSYMQSFQEKVNP 322
QY 72 -----HSEKLGHFHYEDIIICLRLLALNEAEVRAAGRLRLRYLIQDSSI 116
DB 323 FIAQKAGLVLLKEHSIRKDLVLTSPMSSIQNLLNGKLIASAAYRCVRLINSIF 382
QY 117 LQVLLKLVYLIARCDIQOSNEVERTOALRLVRKMITVNASLPSSVTNSLI-AVGND 175

DB 383 IDELLELRDLDAFIILISLAKONSFQIEREQALQWVRFIEYN-----NGVTQIMQAIISC 437
QY 176 GLQERDMVRACIAIIICELALQNPVVALRGGLNTILKNVIDCQLSRINEALITITLHLL 235
DB 438 VEKPEDSLRHMALETLLLELCFVAPENVKECRGMVIEGLQDYTSFSLASVILDTILQLM 497
QY 236 NHPKTRQYVRADVLELERILAPYTDHYRHSPTASQOLKDEAREARFLASKMGIIATFRW 295
DB 498 ATHKTRQHFLEDFNVSVLTTFSDTNTKSNVVERMQ-----NASTL-ISITLSY 547
QY 296 AGIINLCKPGNSIQSLICVLPINMEIRRGLEVLVYDIFRL-PLPVVTEEFIEALLSVD 354
DB 548 NGFMLPSNNFKPLQALVSPFOIP--ICAQYLIDIFLDVLKIKPLPYKPR----- 595
QY 355 PGRFODSWRLSDGFVAEAKTILPHRARSRDPDMONYLALILSAFIRNGLLGLVEVIT- 413
DB 596 -GRHSHPK-----PIPSQYKECMVNRQALALIVLILENSEFVPHLLELLNE 642
QY 414 -NSDDHISVRATILLAGELHMAANTILPHSHSHHLHCLPTLMNMAASFOIPKEKRLASAA 472
DB 643 EDRDDHLVAKGRYLLTEYFNLRMNLDVKYT-----SVSKPIYKENFYVNT 690
QY 473 LNCLEKRFHEMKRGPYSLHLDHIIQKAIATHQKRDQVLRVQKIDIFILKDTTEALLNL 532
DB 691 POFKKIAYKMR-----NENTIGMSGIDYAQNIKSFSKNIKENTLLREVDVDFRFRM 742
QY 533 -RDSOVLOHKNLENNWNLIGTILKWPNNVNRKYD-RQLHRFVRRLLYFYKPSKLYAN 590
DB 743 VYDSKVLQTKDFTRWNNIINELLEGLPLINKQLEELVKSTFIRRLVYFVPLRLFSN 802
QY 591 LDLDPAKAKOLTVVGCQTFELLESERDQGVLEDLVKDIVQ-----WLNASSGMKPERSL 646
DB 803 VNGKAGSQKTVQVQCQFFKLTATPE-GMKILMDTKIIPOLASIMPRAMEGNTSGNIP 861
QY 647 QNGLLTLTSHOYFLIFIGTILSCHPHGVKLEKCSVFQCLLNLCSLKNODHL--LKLTVSS 704
DB 862 NKNKLRKRIIFGYPKFIFIGLQKNGVHILTRWNFFTVIYKMFQESKLGLEFLALLTPE 921
QY 705 LDYSDRGLARVILSKILTAATDACHLYATKHL-----RVLLRANVEFNNWG 751
DB 922 LDLKYSSHCIRVIIGKALVVAKEVRIEATKIHGDKLLESTKESDLKANKVKLOQFK 981
QY 752 IELVLTQLHDKNTTSSALDIDECEDKANLHALIQMKPALSHLGPDKGLL----- 803
DB 982 MEMLTRQLYD-----LSPSVAVADOA-----LYECIVAGNSEELGHSFRMFLNQWFI 1031
QY 804 ---LLRFLSLPKGFSYLNERYGYVAKQLEKWHREYNSKYVDLIEBQL--NEALTTYRKPV 858
DB 1032 RSPILFELLSRPYGFQQLNEINFVKEERDSWLSKKNIEYVHVEBFLKKNESI----- 1084
QY 859 DGDNVRSNORLQRPVYLPILHYQLVHHKHTGCHLLE-----VONIITELCRNVRT 911
DB 1085 ---NAKSITFOQKSR-----LPLHFYESLTKTEDGILLLSQTDGLVTFMNVIKKYVNGNM 1137
QY 912 PDLKWEIEKIKLASLWALGNIGSSNMGLNLQENVIPDLKLAOCOEVLSIRCTCVV 971
DB 1138 ATVENAKELDLKAALWCVGFIGSTELGILLDNLVSDVEIEVAYNASVTSVRTAFV 1197
QY 972 LGLIAKTQGCDDILKCHNWD-AVRHSRKHLVPVDDVEQCLCNELSSIPSTLSLSESTS 1030
DB 1198 LGLISMTREGCEILDEMGWNCVSVQDEPIGIALPNRLDRF-----LSYNEHKWS 1247
QY 1031 SRHNSSESVPSSMFILEDDRFSGSSTSTFPFLDINEDTEPTFYDRSGPIKQNSFPFAS 1090
DB 1248 AFGYSDEMI---VFNKSDG-----DLIEKCLPIEFDLDKLLEKOTAEPLAN 1292
QY 1091 SKLVKNRILNLTLPNKKHRSDDPKGKLSSEKTSNRRIRTLTEPSV-DFNHSDDDFTP 1149
DB 1293 EKIITKNYDNDIT-----SQTITVSGENSLFANEG-----LSSPYVQYRNDSDSIE 1340
QY 1150 ISTVQKTLQLETSFMGN---KHIEDTGST--PSIGENDLKFTKNP 1189

Db 1341 SKVLHIVSQLGNHILSNHAKVEITEINNKYGPRLFENEKMFVKVF 1385

RESULT 14

AAB41158

ID AAB41158 standard; protein; 141 AA.

XX AAB41158;

AC AAB41158;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF922 polypeptide sequence SEQ ID NO:1844.

DE

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; immunostimulant; osteopathic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antitviral; antibacterial; antifungal; antirheumatic; antithyroid; antianemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antinflammatory disease; coagulation; thrombosis; contraceptive.

XX Homo sapiens.

OS

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

PR 02-APR-1999; 99US-0127636P.

PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI: 2000-602362/57.

DR N-PSDB; AAC75367.

XX Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 1428; 5507pp; English.

XX AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antitviral; antifungal; antirheumatic; antithyroid; and antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to

CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 141 AA;

QY Query Match 7.9%; Score 692; DB 3; Length 141;

XX Best Local Similarity 100.0%; Pred. NO. 3.2e-47;

XX Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1366 MKANSFESRLTPSRFMKALSYASLDKEDLLSPINQNTLQRSSSVSRMVSSATYGGSDYI 1425

DB 1 MKANSFESRLTPSRFMKALSYASLDKEDLLSPINQNTLQRSSSVSRMVSSATYGGSDYI 60

QY 1426 GLALPVDINDIFQVKDIPYFQTKNI PHDDRGARAFADHAGGLPSGTGGLVKNSFHLRQ 1485

DB 61 GLALPVDINDIFQVKDIPYFQTKNI PHDDRGARAFADHAGGLPSGTGGLVKNSFHLRQ 120

QY 1486 QMSLTRIMNSIHSDAS 1501

DB 121 QMSLTRIMNSIHSDAS 136

RESULT 15

ABG01059

ID ABG01059 standard; protein; 200 AA.

XX ABG01059;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #1050.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS65246.

XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

XX Claim 20; SEQ ID NO 31418; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (I) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (I). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 200 AA;

Query Match 7.2%; Score 634.5; DB 4; Length 200;
Best Local Similarity 57.0%; Pred. No. 2.6e-42;
Matches 143; Conservative 8; Mismatches 33; Indels 67; Gaps 6;
QY 1291 GSSHTLPRAQSLKAPSIATIKSLADCNFYSSTSDAFGYATLKRLLQQQRMHPSLSHSEA 1350
Db 6 GGKELLPAAGWGEPGVC-----ESSTDLKQC-LK-----AVSHLRS 43
QY 1351 LSPAKDVLFTDTITMKANSFESRLTPGRFMKALSYASLDKEDLLSPINQNTLQSSSVR 1410
Db 44 LEGSTEG---HSC TLPQFSFRDVL---RFMKALSYASLDKEDLLSPINQNTLQSSSVR 96
QY 1411 SMVSSATYGGSDDYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAGGLPS 1470
Db 97 SMVSSATYGGSDDYIGLALPVDINDIFQVKDIPYFQTKNIPPHDDRGARAFADAG---- 152
QY 1471 GTGGLVKNSFHLRQOMSLTEIMNSIHSDSLPLESTDTGLQEHDDNCLYCVCIEILG 1530
Db 153 -----ESTEDTGLQEHDDNCLYCVCIEILG 178
QY 1531 FQPSNQLSAIC 1541
Db 179 FQPSNQLSAIC 189

Search completed: April 9, 2005, 15:12:03
Job time : 101 secs